Proceedings of

Measuring Behavior 2008

6th International Conference on Methods and Techniques in Behavioral Research

Maastricht, The Netherlands
August 26-29, 2008

Editors
Andrew Spink
Mechteld Ballintijn
Natasja Bogers
Fabrizio Grieco
Leanne Loijens
Lucas Noldus
Gonny Smit
Patrick Zimmerman

Noldus
Information Technology

2008

Bibliographical reference

Proceedings of Measuring Behavior 2008, 6th International Conference on Methods and Techniques in Behavioral Research (Maastricht, The Netherlands, August 26-29, 2008)

Edited by A.J. Spink, M.R. Ballintijn, N.D. Bogers, F. Grieco, L.W.S. Loijens, L.P.J.J. Noldus, G. Smit, and P.H. Zimmerman

ISBN 978-90-74821-81-0

P.O. Box 268

6700 AG Wageningen

The Netherlands

www.noldus.com

Printed in the Netherlands

Table of Contents

Keynote Lectures

Symposia

Symposium 1: Analysis of streams and patterns in behavior

Symposium 2: Measuring Behavior using Motion Capture

Symposium 11: Measuring Player Experiences in Digital Games: Behavior, Psychophysiology, and Brain-Based Indicators

Free papers

Free paper session 1: Measuring usability, accessibility and user experience

Free paper session 2: Multimodal analysis of behavior and physiology

Free paper session 3: Latest developments in video and other tracking techniques and automatic behavior recognition

Free paper session 4: Measuring behavior in a clinical context

Free paper session 5: Analysis of patterns of behavior

Free paper session 6: Measuring animal behavior in the field, behavioral ecology

Poster group B

Poster group C

The *Measuring Behavior* Conferences

Measuring Behavior is a unique conference about methods and techniques in behavioral research. While most conferences focus on a specific domain, *Measuring Behavior* creates bridges between disciplines by bringing together people who may otherwise be unlikely to meet each other. At a *Measuring Behavior* meeting, you find yourself among ethologists, behavioral ecologists, neuroscientists, developmental psychologists, psychiatrists, human factors researchers, movement scientists, electronic engineers, software designers, usability specialists… to name but a few. While the research questions and applications are highly diverse, all delegates share an interest in methods, techniques and tools for studying behavior. We have learnt from experience that the focus on methodological and technical themes can lead to a very productive cross-fertilization between research fields. Crossing the boundaries between disciplines and species (from insects to astronauts) can be extremely inspiring. For many delegates, attending a *Measuring Behavior* meeting is an eye-opening experience, to find out which interesting (and often highly relevant) developments are taking place in domains they usually don't venture into.

Measuring Behavior started in 1996 as a workshop in the framework of a European research project "Automatic Recording and Analysis of Behavior", aimed at sharing the results of our project with colleagues from abroad. Organized by Noldus Information Technology and hosted by Utrecht University, *Measuring Behavior* '96 attracted over 150 participants from 25 countries . Encouraged by the international interest, it was decided to make *Measuring Behavior* a recurring conference. In the years that followed, the conference travelled to five other Dutch university towns: Groningen (1998), Nijmegen (2000), Amsterdam (2002), Wageningen (2005) and now Maastricht (2008).

Over the years, *Measuring Behavior* has developed a formula with a mix of ingredients that has proven to be successful. The meeting is always held in a university city where research on human or animal behavior is prominent, with the university hosting the conference and a local professor serving as conference chair (see table below). Noldus Information Technology serves as conference organizer and main sponsor. For a small company like ours, the conference is a major investment. The registration fees just cover the direct expenses associated with the meeting; the hours spent on the organization (several person-years) are on our account. We gladly do this, because we believe that the focused attention on behavior research methods and techniques will eventually lead to a higher demand for our tools. To prevent commercial bias, however, the scientific program is put together under auspices of an independent Scientific Program Committee, consisting of international experts from a broad variety of disciplines (see pages 375-376 for their names). We are very grateful for their effort to review papers and the helpful input during numerous e-mail exchanges.

Year	Venue	Conference chair
1996	Utrecht University	Berry Spruijt
1998	University of Groningen	Jaap Koolhaas
2000	Radboud University Nijmegen	Alexander Cools
2002	Vrije Universiteit Amsterdam	Gerrit van der Veer
2005	Wageningen University	Louise Vet
2008	Maastricht University	Harry Steinbusch

You will notice that the interval between conferences varies between 2 and 3 years. In 2002, in the aftermath of the Internet bubble and 9/11, when travel budgets were cut worldwide, it seemed a wise move to reduce the frequency of the conference. However, we are ready to go back to a biennial rhythm if there is sufficient interest. Also, after six conferences in The Netherlands, the time may have come to go abroad and perhaps even jump across the Atlantic. We would love to hear your view on this. You can let us know when you fill in the conference evaluation form.

In the scientific program, well balanced between human and animal research, one finds a variety of formats for presentation, interaction and exchange of information. The traditional oral papers and poster presentations have always been central to the conference. Increasingly, special symposia – focusing on a current methodological or technical theme – are proposed by experts from various disciplines.

Besides oral presentations in symposia or free paper sessions, the conference program always includes ample time for posters and demonstrations of software or equipment by participants. For several academic inventors, the presentation of their prototype software or hardware tool at *Measuring Behavior* paved the way towards commercialization. This is how CatWalk and FaceReader found their way into the Noldus product portfolio. We hope that scientists will continue to present inventions at *Measuring Behavior* and discuss commercialization opportunities with the vendors present at the meeting.

Another attractive element of the conference is the scientific tours, guided visits to behavioral research facilities and laboratories in and around the hosting university. In 2002, we introduced the tutorials, short courses – mostly about software tools and instruments – taught by expert instructors. This has quickly become a very popular program element. Other program elements are user meetings (organized by manufacturers of research tools), special interest groups, and workshops. Finally, there is the commercial exhibition of scientific books, instruments, and software related to behavioral research.

Over the years, the conference has grown significantly in size, from 153 delegates in 1996 to more than 400 in 2005. At this size, the event is large enough to cover a wide range of topics, yet still small enough for a social program with all delegates. *Measuring Behavior* has also become a truly global meeting: at the previous meeting in Wageningen, delegates came from 40 different countries across six continents.

Now you find yourself at the 6th *Measuring Behavior* conference. The organizers have done their best to prepare an optimal mix of scientific, technical, cultural, social, and culinary ingredients. We hope that you will find *Measuring Behavior 2008* a rewarding experience and wish you a pleasant stay in Maastricht.

Lucas P.J.J. Noldus

Managing Director, Noldus Information Technology BV

L.Noldus@noldus.nl

Measuring Behavior 2008: Highlights of the scientific program

A glance though the scientific program for *Measuring Behavior* 2008 illustrates the extraordinary breadth of disciplines that measure behavior and how that is related to trends in our society. In the aging population of Western nations, health care is a growing area of research, and so it is no surprise to see that we have several symposia on related subjects – animal models for studying depression, measurement of pain, measurement of learning and anxiety, recovery from spinal cord injury and behavior in a clinical context. Usability and other aspects of how people interact with computers and other machines is something that receives increasing attention in society, and indeed has become increasingly important at each *Measuring Behavior* conference. This year we have symposia on measuring usability and user experience, on measuring the experience of people playing digital games, on measuring people in cars and traffic, a workshop on human factors in the cockpit, as well as sessions with a more technical focus such as one about eye tracking measurements. There are more interdisciplinary sessions with a technical emphasis such as one on *Measuring Behavior* with motion capture, a session on techniques for tracking, for *Measuring Behavior* in the field, one on multimodal analysis of behavior and physiology and also a session about telemetry.

A theme that comes back in several sessions this year is about *Measuring Behavior*s in a natural context. This is an area of interest both to people studying animal behavior (explored in the symposium about the home cage as starting point for innovative concepts in behavioral phenotyping) and human behavior (addressed in the symposium on living labs for studying human behavior in the home). The importance of that theme will be debated in a panel discussion (which we can promise will be provocative and interactive) and will also be addressed by two of our keynote speakers. David Wolfer will give a talk on 'Automated testing of mouse behavior in the home cage: can it meet its promises?' and Gregory Abowd will discuss '*Measuring Behavior* in the home environment'.

Our third keynote speaker is Patrick Bateson, and his lecture will be entitled 'What new techniques for *Measuring Behavior* would you die for?' and will focus on future developments in behavioral methodology. His lecture will be immediately followed by a symposium called 'Analysis of streams and patterns of behavior' which follows on from Professor Bateson's talk by addressing fundamental conceptual issues in *Measuring Behavior*.

There are many other subjects covered in the conference this year from statistical analysis to zebrafish and even a symposium entitled 'Measuring the impossible'. As well as the scientific presentations, there are a series of tutorials where lecturers explain the best practices in established techniques for *Measuring Behavior*, and user meetings in which manufacturers present their latest developments and engage in discussion with the audience. There will be something like 250 oral presentations in total and almost 100 poster presentations.

The scientific presentations are all documented in the proceedings book, and after the conference the *Measuring Behavior* web site is converted into an archival site, with abstracts of all presentations, which remain accessible. The web sites of past conferences form a valuable resource on methods and techniques for behavioral research. Finally, selected presentations are published as full papers in the journal Behavior Research Methods, which devotes a special issue to our conference. We are grateful to the Psychonomic Society and to John Krantz, editor, for our pleasant collaboration.

All in all, we can look forward to a fascinating conference with a broad, diverse, and intellectually challenging scientific program!

Andrew Spink

Chair, Scientific Program Committee

A.Spink@Noldus.nl

Conference publications

Measuring Behavior 2008 delegates benefit from a variety of conference publications:

Program book

This book contains a complete overview of all conference events, schedules of sessions and presentations, social events, information about exhibitors, conference organization, and practical information.

Printed proceedings

This year the printed edition of the Conference Proceedings is included in the conference fee! You can find it in your conference bag.

Special issue of Behavior Research Methods

The conference organization has an agreement with Behavior Research Methods (BRM), a peer-reviewed journal (ISSN 0743-3808) published by the American Psychonomic Society (www.psychonomic.org). BRM will devote a special issue to the Measuring Behavior conference. Manuscripts will be subjected to the journal's normal review process. Copies of the special issues can be obtained directly from Psychonomic Society Publications at www.psychonomic.org/brmic/special.htm. The links below take you to the table of contents of the special issues that have been published so far:

- Measuring Behavior '98: www.noldus.com/events/mb98/brmic.htm
- Measuring Behavior 2000: www.noldus.com/events/mb2000/brmic.html
- Measuring Behavior 2002: www.noldus.com/events/mb2002/brmic.html
- Measuring Behavior 2005: www.noldus.com/mb2005/brmic.html

Conference web site

After the conference, the Measuring Behavior conference web site will be converted into an online proceedings site, with abstracts of all presentations, which will remain accessible. The proceedings of past conferences form a valuable resource on methods and techniques for behavioral research. These are links to the proceedings of all five conferences:

- Measuring Behavior 1996: www.noldus.com/events/mb96/mb96.htm
- Measuring Behavior 1998: www.noldus.com/events/mb98/mb98.htm
- Measuring Behavior 2000: www.noldus.com/events/mb2000/
- Measuring Behavior 2002: www.noldus.com/events/mb2002/
- Measuring Behavior 2005: www.noldus.com/mb2005/
- Measuring Behavior 2008: www.noldus.com/mb2008/

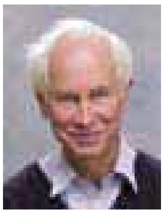
Also keep an eye out for www.measuringbehavior.org - *airs soon!* - for information about future Measuring Behavior conferences!

Keynote Lecture

Professor Sir Patrick Bateson

Sub-Department of Animal Behaviour, University of Cambridge, UK
ppgb@cam.ac.uk

About the speaker



Professor Sir Patrick Bateson, FRS (born 1938) is a biologist and science writer. Bateson is emeritus Professor of Ethology at Cambridge University and president of the Zoological Society of London. He received his BA in zoology and PhD in animal behavior from Cambridge University. Previous academic positions include a Harkness Fellowship at Stanford University and ten years as head of the Cambridge sub-department of Animal Behaviour. He was elected a fellow of the Royal Society in 1983. Bateson is a research scientist and science populariser who has written many books and articles on ethology, animal welfare, measuring behavior, developmental biology and genetics, gives public lectures and broadcasts.

What new techniques for measuring behaviour would you die for?

Many of the classical examples that figured so strongly in the first text books on animal behaviour would not pass editorial scrutiny in the 21st century. Small samples, non-independence of measurements (when measurements were made), pseudo-replication, naïve or improper use of statistics (when statistics were used), lack of adequate controls (when experiments were carried out), not conducting experiments blind and generally using poor experimental design were all flaws in ethological work of that early period.Some of these inadequacies haven't been rectified. Dangers of unconscious data selection remain and not many behavioural biologists run their experiments blind. However, the lecture is intended to be forward-looking. I shall focus on where progress could be made in the efficient measurement of interesting aspects of behaviour without resorting to measurement of the merely trivial. The issues will range from the subtleties of choice experiments, through visualizing data, and discovering context and order in sequential streams of events, to the invention of devices for recognizing patterns in multidimensional arrays.

Keynote Lecture

Professor David P. Wolfer

Institute of Anatomy, University of Zurich and Institute of Human Movement Sciences and Sport, ETH Zurich, Switzerland
dpwolfer@anatom.uzh.ch

About the speaker

David P. Wolfer (born 1960) has been associate professor of anatomy at the Institute of Anatomy at the University of Zurich and at the Institute for Human Movement Sciences at ETH Zurich since 2005. He studied medicine in Zurich and got his doctorate in 1988. His research group investigates the biological basis of cognitive functions, in particular the role of the hippocampus in memory and control of behavior. The group works with genetically modified mice as well as with classical genetics and lesion studies. In the context of this research new methods of behavioral analysis are developed, allowing to measure cognitive functions more efficiently and reliably.

The teaching activities of David P. Wolfer cover the whole microscopic and macroscopic human anatomy with special emphasis on the nervous and musculoskeletal system. At the advanced level, his teaching covers the topics of neuroanatomy and genetics of behavior.

Automated testing of mouse behavior in the home cage:can it meet its promises?

The sequencing of the human and mouse genomes as well as powerful methods to manipulate genes in the mouse have created new challenges for behavioral neuroscientists. The rat as the traditional rodent model of experimental psychology is now often replaced by mutant mouse lines. Many of them have been created to model disorders such as dementia, depression, schizophrenia or autism, hoping that mouse behavior can model the symptoms of human patients. Chemical mutagenesis screens, large scale transgenesis and gene targeting projects, as well as drug screening programs ask for high throughput behavioral analysis of mice. The attempt to respond to these challenges by adapting existing behavioral paradigms has only been partially successful, leading some to consider behavioral analysis as a bottle-neck hindering the progress of neuroscience [1]. Behavioral tests fail to provide the expected sensitivity and specificity. For example, most mouse models of severe human mental retardation lack deficits in cognitive tests or show only very minor changes which can also be found in mouse lines created for other purposes. Behavioral analysis of mice has gained the reputation of being unreliable with results failing to reproduce from test to test and from lab to lab. Finally, because behavioral testing is slow, the generation of mouse models often takes less time than their functional analysis.

Several factors need to be considered. Many of the current tests have originally been developed for rats and were transferred to mice without sufficiently considering the behavioral differences between the two species. Also, many tests are used in a careless way outside the context for which they have originally been validated. In most experiments the mice are socially isolated, handled by humans, transferred into novel and aversive apparatus, and their diurnal activity cycle is disturbed. This introduces stress, increases variability and complicates interpretation. Behavioral testing of mice is slow and labor intensive, because they adapt slowly to experimental procedures and are slow learners in many tasks. In addition, most applications require that animals be tested in a battery of multiple tests. To make things even worse, the behavior of mice tends to be variable and their performance unreliable, creating a need to analyze relatively large samples. In an attempt to reduce costs and to standardize conditions, housing environments are often inappropriate, leading to stimulus deprivation and development of behavioral stereotypies. Finally, to eliminate genetic variability most experiments use inbred strains which have been bred for morphological and/or behavioral extremes.

During the past few years, behavioral testing of mice within their home cage has emerged as a new approach promising to solve many of the problems of conventional behavioral tests [2]. The wish list of features of an ideal system is long. It should minimize the need to handle the animals and allow them to express their behavioral repertoire in an enriched an social environment. The cage should have sensors collecting rich data, permitting not only monitoring of activity but also identification and quantification of defined behaviors including social interactions. Recording of ambient variables (temperature, light, etc.), of body weight, and if needed of physiological parameters should also be possible. As cues, reward or punishment, the cage should permit to present stimuli of various modalities at defined locations and precisely controlled time points. Further, the system should have operant elements permitting animals to interact with the system and to influence its operation. The software should permit flexible and intuitive design and debugging of complex protocols and support fully automated use once protocols are established. The software should not only provide fully automated control of experiments, but also permit real time monitoring of critical variables and produce alarms if an animal fails to perform or is in a critical state. After the experiment the user expects intuitive and flexible software functions for visualization and exploration of the data, data mining tools and the opportunity to fully automate established analyses, including statistics.

Several systems have been developed that implement part of the features in the above wish list. PhenoTyper by Noldus, HomeCageScan by CleverSys, SmartCube by PsychoGenics bring video-tracking technology into the home cage of a single animal and combine this with sophisticated image analysis software to quantify and characterize movements. These systems can be used to monitor spontaneous behavior, but most of them also offer the possibility to present stimuli and to add operant elements. The PhenoMaster project of TSE Systems plans to add interactive elements to extend the possibilities of the existing metabolic cage LabMaster. This system already permits to record physiological parameters and to monitor activity using IR beams or heat sensors. Laboras by Metris places the cage on piezo sensors to record the vibrations generated by the animal and uses specialized software to identify and quantify defined behavioral elements. IntelliCage by NewBehavior uses transponder technology and is the only system to allow monitoring and testing of mice housed in social groups. The IntelliMaze project develops automated gates than can guide mice from the group cage to specialized arenas for additional tests.

The availability of systems permitting automated testing of mouse behavior in the home cage represents an enormous progress. In addition, large efforts are being made to further improve, refine and extend these systems and to increase their usefulness by developing and validating test protocols for various applications. However, many of the actual problems in mouse phenotyping are of conceptual nature and cannot be resolved by technical developments alone. Mere automation of already existing tests will not automatically improve their ethological relevance, nor will it resolve the question whether standardization or controlled variation yield more valid data. Fully automated high throughput screening of mouse behavior may soon be a reality, but is unlikely to be more than a fist step if the goal is to understand behavior [3]. Systems permitting automated testing of mouse behavior in the home cage are a very welcome and powerful new tool that may help us to solve many burning problems of mouse phenotyping, but we first need to learn how to use them in creative ways.

References

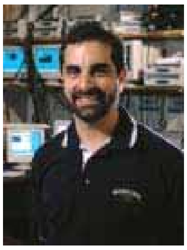
1. Tecott, L.H., and Nestler, E.J. (2004). Neurobehavioral assessment in the information age. *Nat Neurosci* **7**, 462-466.
2. Spruijt, B.M., and De Visser, L. (2006). Advanced behavioral screening: automated home cage ethology. *Drug Discovery Today: Technologies* **3**, 231-237.
3. Crabbe, J.C., and Morris, R.G. (2004). Festina lente: Late-night thoughts on high-throughput screening of mouse behavior. *Nat Neurosci* **7**, 1175-1179.

Keynote Lecture

Professor Gregory D. Abowd

School of Interactive Computing, GVU Center and Health Systems Institute, Georgia Institute of Technology, Atlanta, Georgia, USA
abowd@gatech.edu

About the speaker



Gregory D. Abowd (pronounced AY-bowd) is the Distinguished Professor of Interactive Computing in the College of Computing at Georgia Institute of Technology, Atlanta, USA. His research interests lie in the intersection between Software Engineering and Human-Computer Interaction. Specifically, Dr. Abowd is interested in ubiquitous computing (ubicomp) and the research issues involved in building and evaluating ubicomp applications that impact our everyday lives. In the College of Computing, he is a member of the School of Interactive Computing and the GVU Center and the GeorgiaTech Broadband Institute.

Dr. Abowd directs the Ubiquitous Computing Research Group in the College of Computing and GVU Center. This effort started with the Future Computing Environments research group in 1995, and has since matured into a collection of research groups, including Dr. Abowd's own group. The FCE Group now consists of a federation of many faculty in the College of Computing.

Dr. Abowd received the degree of B.S. in Mathematics and Physics in 1986 from the University of Notre Dame. He then attended the University of Oxford in the United Kingdom on a Rhodes Scholarship, earning the degrees of M.Sc. (1987) and D.Phil. (1991) in Computation from the Programming Research Group in the Computing Laboratory. From 1989-1992 he was a Research Associate/Postdoc with the Human-Computer Interaction Group in the Department of Computer Science at the University of York in England. From 1992-1994, he was a Postdoctoral Research Associate with the Software Engineering Institute and the Computer Science Department at Carnegie Mellon University.

Measuring behavior in the home environment

Since 2000 at Georgia Tech, we have been fortunate to have a unique laboratory environment, the Aware Home (http://www.awarehome.gatech.edu), for the study of technology in the domestic setting. The research agenda at the Aware Home spans many topics, but at the core of our work is the assumption that one day there will be the opportunity and the necessity for home environments to possess a greater awareness of the activities and whereabouts of its inhabitants. In this talk, I will provide an overview of the various research activities that support the agenda of attaining better awareness and representation of behavior in the domestic setting as well as the motivating health applications that leverage this awareness. Those health applications address the general topic of chronic care management, ranging from aging in place to diabetes management to support for developmental disabilities.

Much of the research in activity recognition has used special purpose sensors, ranging from costly optical and acoustical sensors to inexpensive motion and physiological sensors. These solutions have opened up the possibility of creating awareness, but have largely ignored some of the practical issues of widescale deployment. I will present a new idea, *infrastructure mediated sensing,* that leverages the built-in capabilities of a home to provide much of the low-level sensing that enables recognition and measurement of behavior. This idea promises to take activity recognition research in the home to a new level.

Analysis of streams and patterns in behavior Symposium

John C. Fentress
Department of Psychology, Dalhousie University, Halifax and Oregon, jcfenterprises@qwest.net

When we look at behavior we face many challenges. The first is what we watch for, and thus what we record. The human eye, or ear, cannot help but miss much. Our perceptions are limited, and often driven by prejudices we do not perceive.

The question is then what do we do. Sometimes we have to develop technologies to aid us. At other times we simply need to do our best to think more clearly.

The participants in this symposium have each contributed importantly to our understandings of behavior, from complementary but individual perspectives.

The dual themes of this symposium are streams and patterns in behavior, across levels and perspectives of analysis. The term streams is simply a reflection of how ever else we think about behavior, behavior is a dynamic process. The term pattern refers to the fact that we break behavior into units, pieces. We apply nouns, and try to identify.

Patrick Bateson has led multiple aspects of understanding behavior, and will launch our symposium in his keynote address.

Wayne Aldridge and Kent Berridge have made important contributions to see how patterns of behavioral expression are brought into the realm of brain science. Ilan Golani and Yoav Benjamini have pioneered ways to link spatial aspects of behavior and their temporal properties together. Alan Kaluev has expanded our perceptions to show ways we can link movement properties to deep issues such as emotionality. Magnus Magnusson has led the way in pursuing how we link temporal and sequential aspects of behavior together. M. Cabanac and A. J. Cabanac remind us that whatever the value in taking mechanical notes in animals we should not let ourselves ignore possibilities of higher order functions such as potential animal consciousness.

Symposium contents

Stepping outside the traditional “science” box

John C. Fentress
Department of Psychology, Dalhousie University, Halifax and Oregon, jcfenterprises@qwest.net

Ethologists and other students of behavior are faced with a difficult challenge. Their goal is to see how individual components of behavior fit together into higher order patterns. Clearly this provides enormous statistical problems. In this presentation I wish to move beyond these issues of statistics, although they are obviously critical [1].

The more detailed our observations, the more components we can record and the greater number of relationships we need to consider. Clarity in observation is the first challenge. What do we look for, and what do we ignore? Often standardized measures leave out potentially important features. For example, I talked with a professional dancer who asked me how I recorded breathing patterns in the wolves I was observing. I had not thought about doing that. But changes in breathing can signal actions that are not only being performed at the moment, but also actions that are likely to occur next. As illustration, we as other creatures often take a deep breath prior to actions that involve high exertion.

Such comments from my non science friends got me thinking about what we tend to see and record as scientists. As one example, the timing and sequencing of behavior, as they occur together, is for the most part imperfectly analyzed. Our friends in music and dance have much to offer. If we think about music, for example, we deal with notes, rhythms, melodies and so forth. It is often useful to play our behavioral records as if they are natural compositions. We have done that with rodents as well as wolves, and the results are often enlightening. As a single example, think of chords where two or more action properties occur together or in overlapping combinations.

Here I introduce what I think is the most difficult challenge to our measures of behavior. I call it the “pieces – relations” problem. The standard way to think about behavior is to divide an observed stream of events into components, and then ask how these components fit together in time. An obvious example is human speech. Combinations of phonemes produce words which in turn allow speakers to articulate sentences and make points. But the phonemes we articulate are not truly independent of one another. Individual phonemes can be adjusted as a function of other sounds that precede or follow. The “pieces” of our speech are affected by the rules through which they are connected in time.

Similar things happen in animal actions. Individual properties of behavior are adjusted by their neighbors. Thus “pieces” of behavior which we isolate may not be truly “fixed”. They can vary in terms of their broader contexts of expression. I will show how this applies both to rodent grooming and wolf vocalizations. My colleague, Simon Gadbois, has applied issues of prosody in human speech, super-segmental modulations in tone and other features, and shown how they may apply to animal actions. How do we measure these modulations that are superimposed across action components we normally record? [2]

There is a deep issue here. While it is tempting to isolate “pieces” of behavior and then to see how these pieces are strung together, it is easy to forget that these “pieces: can be modulated by their contexts of expression. One image is that of a series of rubber bands that are pulling upon one another, thus changing shape by actions in their neighbors. From a statistical point of view this is a potential nightmare.

The sequencing and timing of behavior are often looked at independently, but these features of behavior are not truly independent. As a simple illustration the duration of face licking movements in rodents is a predictor of which other actions are most likely to follow next. To my knowledge we do not have adequate methods to look at both the timing and sequencing of action properties within a coherent framework.

Those involved in the brain sciences face a similar problem. Our brains are “modular” in important ways. Different areas are primarily involved with different functions. However, these “modules” are not truly independent from one another, and only by examining them in the broader contexts of their operations can one truly achieve an understanding of their operations. Even in invertebrate studies of ethologically derived concepts, such as “central motor programs”, studying systems in their isolation provides a distorted picture of how they operate when connected at the level of the intact organism. Mechanistic and systems views of both brain and behavior provide a conceptual dilemma that we have not successfully come to grips with.

I have found students of music and dance to be especially helpful. In each case they look at the flow of action properties in terms of their relations to one another. As in speech, a dancer or musician blends individual events in terms of the neighborhoods within which they occur. Computer speech and music can sound choppy simply because this blending is not applied.

What does this mean for how we record and interpret behavior? I offer a few thoughts and invite others to share their views.

The question becomes: What is a piece of behavior? We label actions in terms of nouns, but even that linguistic necessity can place a freeze frame on patterns of expression that in reality are much more fluid. .

This presentation seeks to supplement new technologies in scientific study of behavior. There are the most important analytical tools that we have established.. Noldus has been a leader, and in my judgment its contributions are of major significance. What I suggest is that at a conceptual level we have a long way to go, whether looking at the behavior of intact individuals, social groups, or neural mechanisms. Our friends in the arts often have sensitivity to issues that we would do well to hear.

References

1. Martin, P. Bateson, P. (2007). *Measuring Behaviour: An Introductory Guide.* Cambridge University Press, Cambridge.
2. Fentress, J.C., Gadbois, S.(2001). The development of action sequences. *Handbook of Behavioral Neurobiology,* E. Blass, editor. Kluver Academic / Plenum Publishers, New York. 393-431.

Measuring consciousness in animals

M. Cabanac and A. J. Cabanac
Faculty of Medicine, Laval University, Quebec, Qc, Canada, michel.cabanac@phs.ulaval.ca

What is consciousness? As often the case with abstract concepts, most dictionaries provide vague definitions. The Oxford English Dictionary defines it tautologically as "The state of fact of being conscious, as a concomitant of all thoughts, feelings, and volitions," and also as "The totality of the impressions, thoughts, and feelings, which make up a person's conscious being." The Dictionnaire Larousse defines *Conscience* as "The perception, knowledge more or less clear that one can have on self existence and of that of the external world" (our translation). In the following, consciousness will be defined axiomatically as: *An abstract private model of reality—with four dimensions: quality, intensity, hedonicity, and durati*on. The human mind is thus able to call up a broad range of apprehended, recollected, or even totally imagined realities. The result is increasingly complex mental activity: thoughts, feelings, and emotions assume a life of their own within a space that is relatively independent of simple stimulus-response pathways. When this space includes a representation of oneself and how this self interacts with reality, we have the beginnings of self-consciousness.

Consciousness was long considered a human privilege, all other animals being merely machine-like beings [1]. This view was challenged when Darwin [2] pointed out that other mammals could express emotion. The question then faded into the background, largely because of the excesses of psychoanalysis and the efforts of the behaviorist school to make behavior the only object of study, to the exclusion of underlying thought processes [*3*]. Recently, there has been renewal of interest in animal consciousness and a growing acceptance that humans are not the only thinkers. Indeed we do not have direct access to other humans consciouness and we accept indirect evidence for the existence of human consciousness in other people, *i.e*., the verbal and behavioral signs that they provide. We postulate that the same verbal and behavioral signs mean consciousness for ourselves and must be a property all humans share. It is legitimate to question why should similar indirect evidence be rejected when it comes to animals? Although less direct than that provided through verbal communication, such evidence is available [4-8].

In this presentation we shall report the experimental signs of emotion and sensory pleasure that we obtained in animals. However the observer must be prudent and remain permanently aware that the evidence is always indirect [9]. For example many fish complex behaviors such as cheating, altruism, species recognition, individual recognition [10] that we would be tempted to consider signs of consciosness, can be explained on the basis of reflexes. Also, the complex foraging and social communication behavior of bees is often considered intelligent and 'conscious;' however there is evidence that it is purely reflexive [11], in the same way as a computer can be artificially intelligent.

We may accept that self-consciousness is a human property[1] -but we remain with the question of which among the animals are conscious? And which are not? At what point in evolution did the nervous systems cease to operate on a reflexive basis only -when did consciousness arouse [12]? Could/would it be in apes? Mammals? Vertebrates? Invertebrates? Our recordings of emotional fever and tachycardia, and behavioral mimics of sensory pleasure provided signs of consciousness in Mammals, Birds, and Reptiles but not in Amphibians and Fish. Thus, it may be hypothesized that the emergence took place in Vertebrates between Lissamphibia and Amniota, *i.e*., among the amniotes, common ancestors of present-day Mammalia, Chelonia, Lepidosauria, and Archosauria. That conclusion will be supported from several additional lines of 'shareable' experimental evidence obtained from the litterature such as, sleep structure, and play behavior.

References

1. Descartes, R. (1650) Traité des passions de l'âme (Elsevier, Amsterdam).
2. Darwin, C. (1872) The expression of the emotions in man and the animals (University of Chicago Press, Chicago,).
3. Watson, J. B. 1919) Psychology from the Standpoint of a Behaviorist (Lippincott, Philadelphia).
4. Jerison, H. J. (1973) Evolution of the Brain and Intelligence (Academic Press, New York).
5. Burghardt, G. M. (1999) Conceptions of play and the evolution of animal minds Evolution and Cognition **5**, 114-122.
6. Dawkins, M. S. (1993)Through our eyes only? (W. H. Freeman & Co., Oxford).
7. Griffin, D. R. (1992) Animal Minds (University of Chicago Press, Chicago).
8. McFarland, D. J. 1985) Animal Behaviour (Pitman, London).
9. Heyes, C. M. (1994) Theory of mind in non human primates Animal Behaviour **47**, 909-919.
10. Bshary, R., Wickler, W., Fricke, H. (2002) Fish cognition: a primate's eye view Animal Cognition **5**, 1-13.
11. Gould, J. L., Grant-Gould, C. (1995) The Honey Bee (Scientific American Library, New York).
12. Bullock, T. H. (1967) in Neurosciences Research Symposium Summaries F. O. S. e. al., Ed. (MIT Press, Cambridge, Massachusetts), vol. 2, pp. 203-327.

1 *And possibly of some apes.*

Problems of behavior measurements

Ilan Golani[1] and Yoav Benjamini[2]
[1]*Departments of zoology, Tel Aviv University*
[2]*Department of Statistics and Operation Research, Tel Aviv University*

The suggestion that a "good" description of behavior is what a good ethologist considers to be a good description highlights a necessary but not sufficient condition for high-quality descriptions. This is because also good ethologists can err. How, then, if not on the basis of a claim for authority, can one tell the difference between high and low quality descriptions? One help comes from behavior genetics, where differences in behavior are used for the localization of genes on chromosomes. The QTL method used for such analysis depends critically on the quality of the measurements of behavior. If under the same standard conditions one inbred mouse strain is found in one laboratory to be significantly higher on a behavioral measure than another strain, and found to be significantly lower on the same measure in another laboratory, then the results are *not* replicable across laboratories and therefore useless for gene localization studies. Statistical replicability across laboratories thus becomes an objective yardstick for both the relevance of a behavioral measure and for the estimation of the quality of its measurement[1,6].

The high benchmark required for obtaining replicability, demands high quality data. Obtaining such data is, therefore, not a luxury but a constraint dictated by the requirement of replicability. In kinematic studies this implies extensive preparation of the data for analysis, including elaborate smoothing[5] and data segmentation procedures[2] (http://www.tau.ac.il/~ilan99/see/help/), since using first, second, or even third derivative measures like velocity, acceleration, jerk, and curvature drastically increases the system's noise, thus putting severe demands on the quality of these procedures.

The raw data of the movement material collected by tracking systems are kinematic variables such as the time series of location data and their respective calculated derivatives (at the path scale), and movements of the parts of the body (interlimb coordination at the joints scale).

Analysis reveals that these variables sometimes form discrete patterns. Discrete patterns thus constitute the results of the study, not its beginning. If these patterns have a fixed content, then this content can be described only once, for all patterns, and analysis can proceed by using these identical patterns as the building blocks of behavior. However, in the majority of cases the content of these patterns is variable. A premature encapsulation of kinematic features into such patterns, whether by a human observer or by a neural network trained by a human observer, yields "behavior patterns" whose variable content becomes inaccessible for further analysis. These packaged and labeled building blocks may be useful for counting frequencies in time and space, but they constitute "black boxes" as far as a moment-to-moment dynamic analysis is concerned. Hence the numerous ethograms - lists of inert labeled behavior patterns left by classical ethology, which are useful as first-approximation-descriptions but useless as far as the comparative dynamics of behavior are concerned.

In the current computational age, segmentation and packaging of the stream of behavior into discrete patterns is, fortunately reversible and therefore not problematic. Since the time-series of kinematic data are indexed, segmentation is performed at the indexing level, leaving the kinematic time-series intact and accessible for any other type of analysis or any other type of segmentation. Dissecting the flow at the indexing level preserves the transparency of the patterns and allows one to segment the flow in several compatible ways, each highlighting other aspects of the organization of behavior. For example, in our studies of mouse exploratory behavior we segment the path traced by the mouse into lingering episodes and progression segments, based on their speed profile[2]. On the one hand, the topographical distribution of lingering is used to define preferred places[4], and the probability of their performance at particular locations is used to define locational memory[3]. On the other hand, the speed profile within lingering episodes is used to calculate average lingering speed[6] – a highly heritable and discriminative measure characterizing the level of activity during staying-in-place behavior across strains and preparations. By examining the content of lingering episodes at the joints scale, one can also readily establish the momentary level of familiarity the mouse has with the novel environment: in a novel environment mice perform horizontal head scans, whereas in a familiar one they tend to also perform vertical scans. Furthermore, a scan in a particular direction often forecasts the direction the mouse is going to take next. The mouse's location is disclosed at the path scale, whereas the direction of its attention and its intentions (see Fentress, this symposium) are disclosed at the joints scale. Segmentation thus delineates the high-level units, whereas the dynamic content of these units relates them to history or topography, or modulates their significance. The higher level progression segments and lingering episodes can further be assembled into incursions (forays into the center) and excursions (roundtrips from home base), which in turn can be used to define higher level constructs like familiarity[8], locational memory[3], and anxiety[7].

References

1. J. C. Crabbe, D. Wahlsten, B. C. Dudek (1999), *Science* 284-1670.
2. Drai D, Benjamini Y, Golani I. Journal of Neuroscience Methods. 96 (2) (2000)
3. Dvorkin A, Benjamini Y, Golani I (2008) *PLoS Comput Biol* **4(2)**.
4. Golani I, Benjamini Y, Dvorkin A, Lipkind D, Kafkafi N. In: The Behavior of the Laboratory Rat – A Handbook with Tests Edited by Ian Q. Whishaw and Bryan Kolb. Oxford University Press.
5. Hen I, Sakov A, Kafkafi N, Golani I, Benjamini Y. (2004) *Journal of Neuroscience methods* **133** (1-2): 161-172.
6. Kafkafi N, Benjamini Y, Sakov A, Elmer GI, Golani I. (2005) *Proc Natl Acad Sci U S A*. 22;102(12):4619-24..
7. Lipkind D, Sakov A, Kafkafi N, Elmer GI, Benjamini Y, Golani I. (2004). *J Appl Physiol.* **97(1)**:347-59.
8. Tchernichovski, O., Benjamini, Y. (1998) *Biological Cybernetics* **78(6)** 433-440

Understanding brain affective states by measuring animal grooming pattering

Allan V Kalueff and Justin L LaPorte
National Institute of Mental Health, NIMH, NIH, Bethesda, USA

The complex patterning of grooming in rodents, which proceeds in a cephalocaudal direction and involves several distinct stages, can be dissected into its constituent parts and microstructures. Grooming and its patterning have been shown to be sensitive to a number of stressors, making it an ideal target for manipulation in studies of experimental (animal) models of affective disorders like anxiety, depression or obsessive compulsive disorder. Discussed during the talk will be recent studies from this and other labs showing how the investigating of stressed (vs. non-stressed) or pharmacologically treated, as well as genetically different (anxious vs. non-anxious inbred strains; mutant vs. wild type) animals helps to understand brain affective states through measuring the disruption of grooming patterning and its regional distribution. Owing to the ever-increasing number of rodent models that have strong grooming phenotypes, this high-throughput in-depth analysis of grooming behavior is becoming promising for biomedical research on affective disorders.

Observing versus seeing, perception versus detection, and data versus nature

Magnus S. Magnusson
Human Behavior Laboratory, University of Iceland, Reykjavik, Iceland. msm@hi.is

Observing the sun, the moon, and the stars, unaided human observers have seen anything from travelling Gods to holes in a dome. Undeniably, in all science, perception of something must in some sense come first for anyone to bother measuring. The importance of direct observation and perception may, however, be exaggerated. Observing is not always seeing what is, but rather what is not -- is simply imagined. Human everyday perception can thus remind of a Rorschach test.

Theoretically and technically aided observation was required for Manâ€™ss recent detection of many structures now taken for granted in our worldview, including the solar system and the Milky Way, the home of all human observers. -- Some of the elements where visible, but their spatial and temporal relationships too hard to discover without adequate theoretical and technical tools. â€“ The even more recently discovered DNA patterns present all around and within, also quickly became fundamental parts of our worldview. Clearly, being in front of human eyes and being important to humans does not guaranty being seen. Imagination, new concepts, special tools and procedures may be strictly required. Since most human problems as well as their solutions are related to human interactions this justifies the deepest scrutiny.

The discovery of atomic particles and interaction relied heavily on advanced mathematics and technology [1]. Mathematics, now defined by most mathematicians as â€œthe science of patternsâ€□ allows description and detection of ever more patterning in nature, often where none was seen (for example, Fractal, Chaos, and Symmetry/Group mathematics). Modern physics thus exemplifies a formalized pattern-view of nature minimally based on unaided perception, but of daunting importance for all modern life. No human interaction patterns can, however, be predicted (as opposite to interpreted or explained) on the basis of particle physics or DNA patterns alone.

Regarding measurement, obviously, phenomena in nature that cannot be detected, cannot be counted, classified, or analyzed in any way. Measuring the size or frequency of invisible patterns such as, for example, polygons, circles, or any other forms of interest in nature therefore presupposes their detection, which again presupposes structural concepts such as â€œpolygonâ€□ and â€œcircleâ€□. Invisible behavioral patterns may thus become â€œvisibleâ€□ given adequate structural hypotheses concerning relationships between their visible parts.

Discovery of behavioral â€œparticlesâ€□ and pattern now depends less on direct perception as video freezes â€œtimeâ€□ and â€œspaceâ€□, while mathematics and computers allow domain-specific analysis that gradually replaces domain-independent statistical analysis.

For example, rather than relying on direct perception of facial behavior, the basic measurements may be automatically recognized changing contrasts in human faces, which then allow the detection of complex and sometimes invisible spatial and temporal patterns. And just like nuclear particles and the solar system some may never be really â€œseenâ€□.

Humans had been talking for numerous millennia before even imagining grammars and only since about 50 years has knowledge of nonverbal interactions benefitted from systematic studies, which are still hampered by lack of adequate models and tools. Discovering the real-time, multi-level, partly parallel structure of everyday interactions thus remains a formidable challenge and the attention shifts to â€œsequential analysisâ€□ and questions about what comes â€œnextâ€□. While frequent behaviors like eye blinks offer a safe guess, but typically of little value for interpretation and/or prediction. For this, reference is usually required to higher order spatial and temporal envelopes (patterns, context).

The t-pattern, t-packet, and derived concepts [2, 3, 4] are attempts at capturing some features of the spatial and temporal â€œprobability landscapeâ€□ of behavioral events. T-packets are at once sequential and non-sequential recurrent structures based on the t-pattern type, which has already allowed the detection of complex patterning in many kinds of behavior and interactions varying from interactions in brain cell networks to childrenâ€™s object exchange and problem solving dyads where little or no structure was found through either direct observation or standard statistical approaches [5, 6, 7].

These hypotheses and illustrative results obtained with the specially designed algorithms are presented. The examples rely on direct perception to a highly different degree, from measures of physical movements of parts of the human face that could be fairly easily automated, to coded acts, that is, directly perceived and classified behavioral entities requiring complex cognitive processing.

Also presented are, so called, â€œGhost cyclesâ€□ (t-ghostcycles), which are cyclically recurring t-patterns of elements each of which does not per se share the cyclical occurrence of the pattern and thus no provide any explanation of the cyclic occurrence of the pattern [8, 9].

Einstein, Max Planck, and Newton mostly analyzed data collected by others, but their work would hardly be labeled simply â€œdata analysisâ€□, something they could simply have delegated to any mathematician or statistician with little or no knowledge of physics or astronomy. In the behavioral sciences, however, analyzing data and analyzing behavior often gets confused. Thus sometimes practically all the (computational) behavior analysis in a project is labeled simply as â€œdata analysisâ€□ or â€œstatistical analysisâ€□ and is happily delegated to, for example, any statistician around. Possibly due to the spectacular progress in computational methodology (including A.I.) wile computer illiteracy is still common, there seems to be much confusion of a) domain-independent standard statistical data analysis requiring no knowledge of behavior and b) domain-specific theory driven research automation requiring deep knowledge of behavior. -- Could this be simply because both cases involve computers and statistics?

References

1. B. A. Schumm. (2004) Deep Down Things: The Breathtaking Beauty of Particle Physics. London: The John Hopkins University Press

2. M. S. Magnusson. (2000) Discovering Hidden Time Patterns in Behavior: T-Patterns and their Detection. *Behavior Research Methods, Instruments and Computers* **32(1)**: pp. 93-110.
3. M. S. Magnusson, Understanding Social Interaction: Discovering Hidden Structure with Model and Algorithms. In Anolli, L., Duncan, S. , Giuseppe, R. & Magnusson, M.S. (Eds). (2005) The Hidden Structure of Interaction: From Neurons to Culture Patterns. Amsterdam: IOS Press.
4. M. S. Magnusson. (2005) Repeated Patterns in Behavior and Other Biological Phenomena. In Evolution of Communication Systems : A Comparative Approach (Vienna Series in Theoretical Biology). D. Kimbrough Oller (Editor), Ulrike Griebel (Editor). London: The MIT Press.
5. Anolli, L., Duncan, S. , Giuseppe, R. & Magnusson, M.S. (Eds). (2005) The Hidden Structure of Interaction: From Neurons to Culture Patterns. Amsterdam: IOS Press.
6. M. S. Magnusson, J. Beaudichon. (1997) "DÃ©tection de "marqueurs" dans la communication rÃ©fÃ©rentielle entre enfants." In J. Bernicot, Caron-Pargue, J. , Trognon, A. (Ed.), Conversation, Interaction et Fonctionnement Cognitif (pp. 315-335). Nancy: Presse Universitaire de Nancy.
7. A. U. Nicol, K. M. Kendrick & M. S. Magnusson. (2005) Communication within a neural network. In L Anolli, S Duncan Jr, MS Magnusson & G Riva, The hidden structure of interaction: from neurons to culture patterns, IOS Press, Amsterdam, NL, pp. 74-78.
8. M. S. Magnusson. (1989) "Structures syntaxiques et rythmes comportementaux: sur la dÃ©tection de rythmes cachÃ©s." *Siences et Techniques de l'Animal du Laboratoire* **14**, 143-147.
9. Magnusson, M.S. (2006) Structure and Communication in Interaction. In G. Riva, M.T. Anguera, B.K. Wiederhold, F. Mantovani (eds.) 2006. From Communication to Presence: Cognition, Emotions and Culture Towards the Ultimate Communicative Experience. Amsterdam: IOS Press.

Brain systems for action sequences

J.W. Aldridge[1,2] and K.C. Berridge[2]
[1]*Department of Neurology, University of Michigan, Ann Arbor, Michigan, United States, jwaynea@umich.edu*
[2]*Department of Psychology, University of Michigan, Ann Arbor, Michigan, United State,. berridge@umich.edu*

Movement, language, and thought occur in streams of complex sequential patterns, a shared feature that has been called action syntax. Instinctive behavior of animals such as natural rodent grooming can be particularly useful for studying sequential patterns of behavior, due to its rich array of complex but predictable movement sequences, with features of action syntax. Rodent grooming has syntactic (rule-driven) sequences with holistic patterns of serial structure [1].

Evidence indicates that brain circuits containing the striatum and substantia nigra causally implement the syntactic sequence, and also code its serial organization in neuronal firing patterns [2;3]. Brain structures such as neostriatum and substantia nigra similarly help control the sequential pattern of normal human action, language and thought, and nigrostriatal dysfunction is linked to human disorders including obsessive-compulsive disorder, Tourette's syndrome, and Parkinson's disease. Indeed it has been suggested that sequencing functions of the striatum in human language/thought may derive from its evolutionary 'action syntax' role in sequencing instinctive actions, including rodent grooming. Our studies aim to clarify how brain systems carry out action syntax functions.

The syntactic behavior during grooming in rats is used to study the role of the basal ganglia in such natural sequential behaviors. Learning more about how neurons code sequential movement may have important implications for treatment and understanding of Parkinson's disease. Our research involves studies of neuronal activity in the basal ganglia. There is evidence that these regions play a role in controlling movement, as well as reward learning and interval timing. Our long-term goal is to understand how individual neurons and neuronal circuits in the basal ganglia might be contributing and processing information related to these processes. We evaluate movements in both normal states and in animal models that simulate human movement disorders. Our principal method is to record electrical activity of individual nerve cells while animals execute either instinctive movements or learned movements in response to sensory cues. From this information, we study the computational properties of neuronal networks activated during our experiments. In these experiments we also activate neural systems by the application of dopaminergic drugs that are known to affect behavior. Currently, our efforts are directed toward examining neuronal mechanisms related to sequences of grooming movements [4], reward learning [5], and the effects of dopaminergic manipulations [5-7]. Also we have studied how excessively rigid behavioral patterns can be caused, in mutant mouse and related models of obsessive compulsive disorder and Tourette syndrome [8]. Our goal is to better understand the brain link that connects animal instinctive actions to human language and thought.

Finally, we are also exploring the relation between action and emotion and motivation, including hedonic impact of sensory pleasures, in striatal processing. One interesting recent finding from our lab is that one region of the basal ganglia, the ventral pallidum, seems to have neural activity that especially represents things that taste "pleasant" or "good" [5;9;10].

Results in our studies indicate show that the brain neostriatum controls the sequential pattern of instinctive grooming behaviors by making action syntax patterns stronger. They also suggest that neostriatum roles in action syntax and patterning behavior may related in some ways to its roles in psychological functions involving emotion and motivation.

References

1. Berridge, K.C., Fentress,J.C., and Parr,H., (1987) Natural syntax rules control action sequence of rats, *Behav. Brain Res*. **23** 59-68.
2. Aldridge, J.W. and Berridge,K.C., (1998) Coding of serial order by neostriatal neurons: a "natural action" approach to movement sequence, *J. Neurosci.* **18** 2777-2787.
3. Meyer-Luehmann, M., Thompson,J.F., Berridge,K.C., and Aldridge,J.W., (2002) Substantia nigra pars reticulata neurons code initiation of a serial pattern: implications for natural action sequences and sequential disorders, *Eur. J. Neurosci*. **16** 1599-1608.
4. Aldridge, J.W., Berridge,K.C., and Rosen,A.R., (2004) Basal ganglia neural mechanisms of natural movement sequences, Can. *J. Physiol Pharmacol.* **82** 732-739.
5. Tindell, A.J., Berridge,K.C., and Aldridge,J.W., (2004) Ventral pallidal representation of pavlovian cues and reward: population and rate codes, *J. Neurosci.* **24** 1058-1069.
6. Berridge, K.C. and Aldridge,J.W., (2000) Super-stereotypy I: Enhancement of a complex movement sequence by systemic dopamine D1 agonists, *Synapse* **37** 194-204.
7. Berridge, K.C. and Aldridge,J.W., (2000) Super-stereotypy II: Enhancement of a complex movement sequence by intraventricular dopamine D1 agonists, *Synapse* **37** 205-215.
8. Berridge, K.C., Aldridge,J.W., Houchard,K.R., and Zhuang,X., (2005) Sequential super-stereotypy of an instinctive fixed action pattern in hyper-dopaminergic mutant mice: a model of obsessive compulsive disorder and Tourette's, BMC. Biol. 3 4.
9. Tindell, A.J., Smith,K.S., Pecina,S., Berridge,K.C., and Aldridge,J.W., (2006) Ventral pallidum firing codes hedonic reward: when a bad taste turns good, *J. Neurophysiol*. **96** 2399-2409.
10. Tindell, A.J., Berridge,K.C., Zhang,J., Pecina,S., and Aldridge,J.W., (2005) Ventral pallidal neurons code incentive motivation: amplification by mesolimbic sensitization and amphetamine, *Eur. J. Neurosci.* **22** 2617-2634.

Measuring Behavior using Motion Capture Symposium

Wim Fikkert, Herman van der Kooij, Zsofia Ruttkay, and Herwin van Welbergen
University of Twente, Enschede, The Netherlands, mb@ewi.utwente.nl

Introduction

Motion capture systems, using optical, magnetic or mechanical sensors are now widely used to record human motion. Motion capture provides us with precise measurements of human motion at a very high recording frequency and accuracy, resulting in a massive amount of movement data on several joints of the body or markers of the face. But how do we make sure that we record the right things? And how can we correctly interpret the recorded data? In this multi-disciplinary symposium, speakers from the field of biomechanics, computer animation, human computer interaction and behavior science come together to discus their methods to both record motion and to extract useful properties from the data. In these fields, the construction of human movement models from motion capture data is the focal point, although the application of such models differs per field. Such models can be used to generate and evaluate highly adaptable and believable animation on virtual characters in computer animation, to explore the details of gesture interaction in Human Computer Interaction applications, to identify patterns related to affective states or to find biomechanical properties of human movement.

Goals to be achieved

- Foster cross-disciplinary knowledge exchange on methods to construct motion models from motion capture data
- Discus and experience (by an industry demo) the state of the art of motion capture systems
- Give a broad overview of the applications of motion capture
- Improve the participants knowledge and skills of the technological issues that are inherently related to motion capture

Topics

The topics covered in the talks are related to the stages and applications of measuring behavior by motion capture technology, such as:

- Smoothing and cleaning the data, e.g. to eliminate noise, to correct for lost markers etc., without loosing valuable details or modifying the empirical data otherwise
- Automated segmentation human motion sequences into gesture units and the recognition of gestures
- Deriving biomechanical and physical characteristics of the person based on analysis of his (dedicated) motion samples
- Deriving motion behavioral characteristics (like smoothness, velocity profiles)
- Deriving both the cognitive and the emotional state characteristics through motion analysis
- Creating a motion model based on captured samples
- Evaluation of the believability of animation generated my a motion model by comparing it with similar captured motion
- Learning motion sequences automatically
- Exploring the effect of prosthesis and other artificial items

Symposium contents

6 DOF Motion Analysis Using Inertial Sensors

Daniel Roetenberg, Henk Luinge, and Per Slycke

Hip compression force estimation with a comprehensive musculo-skeletal model

H.F.J.M. Koopman and M.D. Klein Horsman

Ambulatory estimation of ankle and foot dynamics and center of mass movement

H. Martin Schepers, Bart F.J.M Koopman, Edwin H.F. van Asseldonk, Jaap H. Buurke, and Peter H. Veltink

4 years of FreeMotion: towards practical large scale application of ambulatory 3D analysis of human movement

Chris T.M. Baten

Analysis of human navigation and manipulation motions

A. Egges

Combining manipulation and navigation in virtual environments

B.J.H. van Basten

Using motion capture data to generate and evaluate motion models for real-time computer animation

H. van Welbergen

Using motion capture to recognize affective states in humans

Nadia Bianchi-Berthouze

Online Segmentation of Continuous Gesturing in Interfaces

F.W. Fikkert, P.E. van der Vet, and A. Nijholt

The influence of gender stereotype priming on social action

E. Ngubia Kuria, Luisa Sartori, Castiello Umberto, and Raffaella I. Rumiati

6 DOF Motion Analysis Using Inertial Sensors

Daniel Roetenberg, Henk Luinge and Per Slycke

Xsens Technologies B.V., Enschede, The Netherlands, daniel.roetenberg@xsens.com

Introduction

The use of miniature inertial sensors has become a common practice in ambulatory motion analysis. For accurate and drift free orientation estimation several methods have been reported combining the signals from 3D gyroscopes, accelerometers and magnetometers. Accelerometers are used to determine the direction of the local vertical by sensing acceleration due to gravity. Magnetic sensors provide stability in the horizontal plane by sensing the direction of the earth magnetic field like a compass. Data from these complementary sensors can be used to eliminate drift by continuous correction of the orientation obtained by integrating rate sensor data [1].

By using the calculated orientations of individual body segments and the knowledge about the segment lengths, rotations between segments can be estimated and a position of the segments can be derived under strict assumptions of a linked kinematic chain [2]. This method assumes an articulated rigid body in which the joints only have rotational degrees of freedom. However, a human body and its joints cannot be modeled as a pure kinematic chain with well-defined joints such as hinge-joints and ball-and-socket-joints. Each human joint allows some laxity in all directions (both translational and rotational) other than its main direction of movement [3]. Moreover, to be able to track complex human joints and non-rigid body parts accurately, more than three degrees of freedom, as given by an orientation measurement, are required. Furthermore, importantly, with only orientation driven motion capture, it is not possible to analyze the clearance of both feet, which occurs during running or jumping. Using this approach, it is also not possible to accurately determine the displacement of the body with respect to a coordinate system not fixed to the body.

To provide full six-degree-of-freedom tracking of body segments with connected inertial sensor modules, each body segment's orientation and position can be estimated by, respectively, integrating the gyroscope data and double integrating the accelerometer data in time. However, due to the inherent integration drift, these uncorrected estimates are only accurate within a few seconds.

In this study, a new method is presented to estimate body segment orientation and position by integration of gyroscope and accelerometer signals which are continuously updated by using a biomechanical model of the human body. By facilitating the constraints of the model, notably, the segments are connected by joints, the kinematics of body segments are corrected for drift and other errors. This paper will focus on the joint position update with an example of gait.

Methods

Sensor fusion scheme

The biomechanical model includes the assumption that two body segments are on average connected but with a statistical uncertainty. For specific joints, rotational characteristics can also be described in statistical terms. For example, in the knee, the main axis of rotation is flexion and extension whereas endo rotation and abduction are usually limited to a few degrees and thus statistically more unlikely. Since the sensor signals and the biomechanical model can be described in a stochastic manner, it can be incorporated in a sensor fusion scheme with a prediction and correction step (Figure 1). In the prediction step, all sensor signals are processed using inertial navigation system (INS) algorithms. This is followed by the prediction of the segment kinematics using a known sensor to body alignment and a model of the body. Over time, integration of inertial sensor data leads to errors and increased position uncertainty due to presence of sensor noise, offsets, orientation errors or other errors, e.g. soft tissue artifacts. The correction step includes joint updates, detection of contact points of the body with an external world which constraints the global position and velocity, and optionally, other aiding sensors. Estimated kinematics are fed back to the prediction step to be used in the next frame.

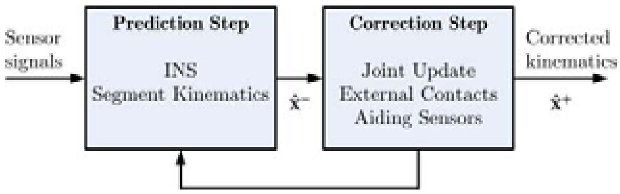


Figure 1. *Sensor fusion scheme. In the prediction step, sensor kinematics are calculated using inertial navigation algorithms (INS) from measured accelerations and angular velocities. Using the biomechanical model, the sensor kinematics are translated to body segment kinematics. In the correction step, joint updates are applied to the segments, followed by the detection of contacts of body points with the external world and optionally aiding sensors.*

Joint update

For each joint, the position relation can be expressed as a linearized function:

$$\mathbf{y}_t = \mathbf{C}\mathbf{x}_t + \mathbf{w}_t$$

where is **x** the state vector at time *t* containing the positions of the two segments, **C** is the measurement matrix relating the state vector to the measurement **y**, **w** is the measurement noise. When two segments are connected, measurement matrix **C** is given by:

$$\mathbf{C} = [\mathbf{I}_3 \;\; -\mathbf{I}_3]$$

$\mathbf{I}_3$ symbolizes the 3 by 3 identity matrix. A Kalman filter is used to estimate the state using the joint relation and the state prediction by the segment kinematic integration step:

$$\mathbf{x}_t^+ = \mathbf{x}_t^- + \mathbf{K}(\mathbf{y}_t - \mathbf{C}\mathbf{x}_t^-)$$

where $\mathbf{x}_t^-$ and $\mathbf{x}_t^+$ are the states before and after the Kalman update, respectively, and **K** is the Kalman gain. The Kalman gain is computed based on stochastic parameters about positional and rotational characteristics for each joint and propagation of errors by each integration step based on the sensor noise. With the Kalman filter update, the kinematics are corrected for drift and the uncertainty of the joint position is reduced (Figure 2).

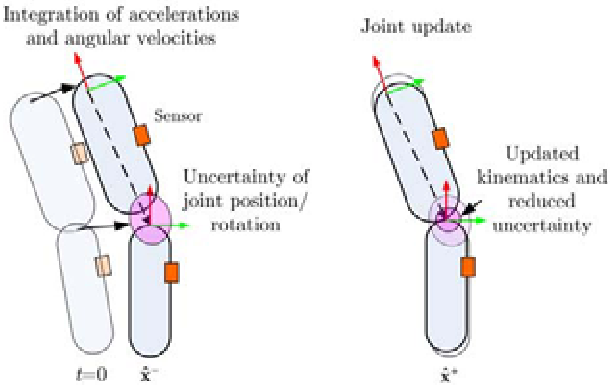


Figure 2. *Integration of accelerations leads to an increased uncertainty about the joint position. After a joint update, the kinematics are corrected and the uncertainty is reduced.*

Experimental set-up

For the experiments, a Moven motion capture system was used consisting of 16 MTx sensors with two Xbus Masters running at 100 Hz [4]. The MTx is an inertial and magnetic measurement unit and comprises 3D gyroscopes, 3D accelerometers and 3D magnetometers (38×53×21 mm, 30 g). For the analysis of the data, the sensors on the feet, lower legs, upper legs and pelvis were used. No magnetometer signals were used.

One healthy subject was asked to walk at a comfortable speed for 10 meters in an office corridor. Data was processed using a translational laxity of the hip joint with a SD of 1cm.

Results

In Figure 3, the kinematics of 2 complete gait cycles are presented. The upper graph shows the position of the hip joint in the global reference frame. The accuracy in position of the foot was 0.19 m after 10 meters. Walking speed was 1.2 m/s. The second graph shows the hip flexion/ extension angle. The dotted lines indicate the standard deviation band of a reference hip angle database of normal walking [5]. The vertical black dashed lines indicate the heel strike events, the red dotted lines the toe off events. These gait cycle events were detected using kinematic properties of the foot, such as height, velocity and acceleration. The third and lower graph shows the estimated movement of the head of the femur within the hip joint of the joint by using the joint update.

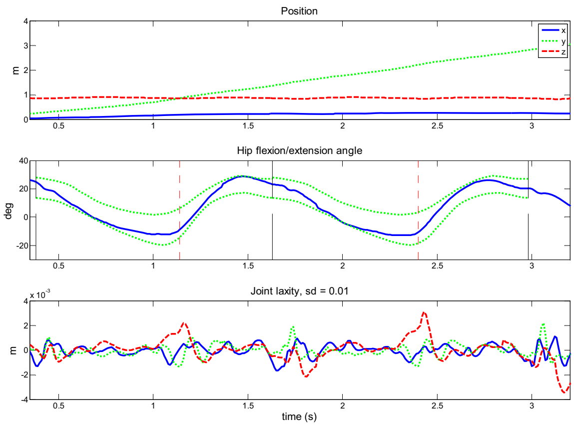


Figure 3. *Kinematics of 2 gait cycles.*

Discussion

The results show the feasibility of using inertial sensors in combination with joint updates to estimate drift-free 3D kinematics. The translation accuracy is within 2% of the travelled distance. Hip flexion and extension angles show a pattern which is similar to a reference gait database.

The observed movement within the joint is mostly due to soft tissue artifacts of the upper leg. As a result, position and orientation changes in the segments around the joint can be measured which are biomechanically unlikely. In the joint update step, the allowed laxity can be set according to rotational and translational parameters as known from literature. Measured motion beyond these values is most likely due to soft tissue artifacts. Therefore, these errors can be reduced.

In processing the data, no magnetometers were used. For longer trials, magnetometers are necessary to provide stability for rotations about the vertical. However, metallic objects can locally disturb the earth magnetic field introducing errors. By using a disturbance model, similar to [1] and rotational constraints in the joint update step, these errors can be minimized.

Since the inertial data can be used accurately for position changes, the presented method allows for dynamic motion analysis including clearance of both feet.

References

1. D. Roetenberg, et al. (2005) Compensation of magnetic disturbances improves inertial and magnetic sensing of human body segment orientation, IEEE *Trans. Neural Sys and Rehab Eng*. **13(3),** 395-405
2. E. Bachmann. (2000) Inertial and magnetic tracking of limb segment orientation for inserting humans into synthetic environments, PhD Thesis, Naval Postgraduate School
3. V. Zatsiorsky. (1998) Kinematics of Human Motion
4. Xsens Technologies, http://www.xsens.com.
5. H.F.J.M Koopman. (1989) The three-dimensional analysis and prediction of walking. PhD thesis, University of Twente

Hip compression force estimation with a comprehensive musculo-skeletal model

H.F.J.M. Koopman[1] and M.D. Klein Horsman[1,2]

[1]Laboratory for Biomedical Engineering, University of Twente, Enschede, the Netherlands, h.f.j.m.koopman@utwente.nl

[2]Human Performance Lab, Faculty of Kinesiology, University of Calgary, Canada, kleinhorsman@kin.ucalgary.ca

Introduction

Muscle forces exerted during human movement provide insight in tissue load and muscle function and/or malfunction. The use of musculoskeletal models in the estimation of muscle force has been widely reported in the literature (see for review [1]). Although model-based estimation of muscle force showed to have clinical potential, many challenges still have to be overcome.

(1) In many models muscle force can instantaneously drop to zero or rise to maximal force (e.g. [2]) However, in reality muscular dynamics will prevent such fast transitions in force. (2) Objective functions based on mechanical measures such as muscle force or stress are frequently used (e.g. [3]). However, the validity of these functions is unknown. A recent study posed an energy-related cost function that had a better correspondence with muscle energy consumption than cost functions based on muscle stress [4]. (3) Since the objective function and its boundaries are a function of anatomical parameters, the outcome of a muscle force optimization is highly dependent on such data. Frequently used datasets are either incomplete or constructed by merging of datasets based on different individuals. This might result in possible inaccuracies caused by inter-individual anatomical differences and co-variance between parameters.

We have developed a musculoskeletal model of the lower extremity based on a recently collected, extensive and consistent anatomical dataset [5]. In the model muscle forces are optimized using a recently developed inverse-forward dynamic optimization (IFDO) method [6], which takes muscle dynamics into account. A recently proposed energy related cost function is used that showed to have a good correspondence with energy consumption [4].

The aim of this study is to evaluate the effect of IFDO and the cost function on hip reaction forces determined with the model, by comparing with existing methods.

Methods

In this two-legged model, 10 joints are crossed with 264 Hill-type muscle elements, defined by muscle parameters such as optimal fiber length. 'Via' points or wrapping geometries were defined in case of a curvature in muscle line of action [5]. Muscle dynamics are described by a third-order muscle model with excitation, active state and contractile element length as state variables. To evaluate the dynamic model properties, gait kinetics was collected for a healthy male subject (age 21, mass 85 kg, length 1.85m). The size of the model and the muscle parameters (e.g. optimal fiber length) were scaled to the length of the segments of the subject. PCSA was scaled to subject mass. The objective function represents the muscle energy consumption of the two-major energy processes in the muscle. The first is the detachment of cross bridges which depends on fiber length and the muscle force. The second process is the re-uptake of calcium which depends on muscle mass and the ratio of actual muscle force and maximal force at a certain length. IFDO was used to estimate muscle forces during gait. This method, to solve the load sharing problem, accounts for muscular dynamics. For each time step at a given state, the minimal and maximal possible muscle force was determined, resulting in respectively a lower and upper limit. These boundaries were used as a constraint, preventing instantaneous drop to zero or increase to maximum in optimized muscle force. After the muscle forces are estimated, an inverse model is used to update the states of the muscle model. In order to satisfy the equations of motion, a second constraint enforces the contribution of the muscles to the joint moment to equilibrate the calculated joint moment. The joint forces are estimated as the combined result from all muscles crossing the joint.

Results and discussion

In agreement with previous model simulations [7], variation in mechanical based cost functions had a small effect on hip compression force. However, in addition, our simulations showed that when the energy related cost function was used instead of a mechanical based function, hip compression force increased with 30%. An optimization with dynamic muscle force constraints resulted in 70% increase in compression force in the hip when compared to an optimization with static constraints. This is in contrast with a previous study showing that static and dynamic optimization solutions are practically equivalent [8].

The current model is based on accurate and consistent anatomical data, which will likely improve the outcome of the model. In this study is showed that besides this effect, IFDO and the energy cost function have a substantial influence on predicted hip forces. It is expected that this will improve the accuracy of joint load estimation in comparison to commonly used methods that exclude muscle dynamic properties as described earlier. However, since the actual hip compression force of the subject in this study was not measured directly, this remains speculative.

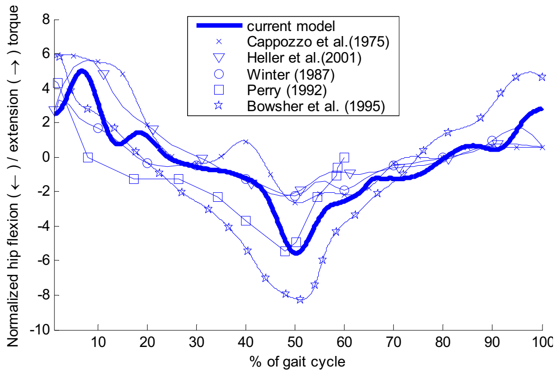


Figure 1. *Hip flexion moment determined with the model normalized to subject length and mass as a function of gait cycle in comparison with normalized moments from the literature*

In the literature a wide range of hip compression forces is reported as a result of differences between subjects e.g. walking speed, style) and used methodology (e.g. # optimized DOF, optimization method, cost function). In general, calculated joint moments, which directly effect hip

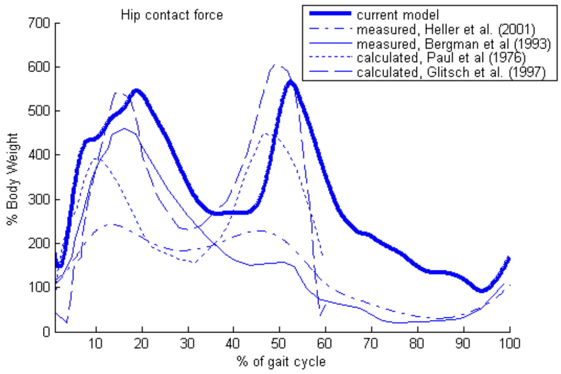


Figure 2. *Hip compression force (% BW) determined with the model as function of gait cycle in comparison with model simulations and in vivo measured compression forces from the literature.*

compression force, fall reasonably well within the wide range found in the literature (See as example hip flexion torque in figure). Despite normalization to subject length and weight, a large variation in amplitude was found between different studies. Such variations can be mainly attributed to differences in walking speed and style for example due to age. The hip compression force determined with the model using IFDO and the energy related cost function was around 2 times larger as measured in an in vivo study [9] (figure 2). This difference is in agreement with the large difference in hip moments between these studies (1.9 vs. 5.5 normalized peak hip flexion moment as shown in figure 1) as a result of lower walking speed (1.08 instead of 1.51 m/s) and subject condition (61 year old with a hip prosthesis instead of healthy 21 year old subject).

Conclusions

Besides the effect of anatomical data on model output [5,6], this study shows that the estimation of hip compression force is highly dependent on the used optimization method and cost function. This emphasizes the relevance of the use of accurate cost functions and optimization methods in order to estimate accurate muscle forces and joint loads.

References

1. Erdemir, A., McLean, S.,Herzog, W., Van den Bogert, A.J. (2007). Model-based estimation of muscle forces exerted during movements. *Clinical Biomechanics* **22**, 131-54.
2. Pedersen, D.R., Brand, R.A., Davy, D.T. (1997). Pelvic muscle and acetabular contact forces during gait. J Biomecanics, 30, 959-65.
3. Heller, M.O., Bergmann, G., Deuretzbacher, G., Durselen, L., Pohl, M., Claes, L., Haas, N.P., Duda, G.N. (2001). Musculo-skeletal loading conditions at the hip during walking and stair climbing. *J Biomechanics* **34**, 883-93.
4. Praagman, M., Chadwick, E.K., Van der Helm, F.C., Veeger, H.E. (2006). The relationship between two different mechanical cost functions and muscle oxygen consumption. *J Biomechanics* **39**, 758-65.
5. Klein Horsman, M.D., Koopman, H.F., Van der Helm, F.C., Prose, L.P., Veeger, H.E. (2007). Morphological muscle and joint parameters for musculoskeletal modelling of the lower extremity. *Clinical Biomechanics* **22**, 239-47.
6. Van der Kooij, H., Van der Helm, F.C. (2003). Human gait analysis application of a new inverse/forward dynamic optimization (IFDO) method to solve the load sharing problem. International symposium on computer simulation in biomechanics. *Proceedings of ISCSB, Sydney, Australia, 2003*.
7. Glitsch, U., Baumann, W. (1997). The three-dimensional determination of internal loads in the lower extremity. *J Biomechanics* **30**, 1123-31.
8. Anderson, F.C., Pandy, M.G. (2001). Static and dynamic optimization solutions for gait are practically equivalent. *J Biomech* **34**, 153-61.
9. Bergmann, G., Deuretzbacher, G., Heller, M., Graichen, F., Rohlmann, A., Strauss, J., Duda, G. N. (2001). Hip contact forces and gait patterns from routine activities. *J Biomech* **34**, 859-71.

Ambulatory estimation of ankle and foot dynamics and center of mass movement

H. Martin Schepers[1], Bart F.J.M Koopman[1], Edwin H.F. van Asseldonk[1], Jaap H. Buurke[2], and Peter H. Veltink[1]
[1]*Institute for BioMedical Technology (BMTI), University of Twente, Enschede, The Netherlands*
[2]*Roessingh Research and Development (RRD), Enschede, The Netherlands,*
h.m.schepers@ewi.utwente.nl

Introduction

Traditionally, human body movement analysis is done in so-called 'gait laboratories', equipped with several measurement systems such as optical position measurement systems, EMG or force plates. The measured signals are used to estimate important gait variables. An important variable is the center of mass, an imaginary point at which the total body mass can be assumed to be concentrated. Several methods exist for center of mass estimation, of which the segmental kinematics method and the double integration of ground reaction force method are the most important ones. Other important variables are joint moment and powers. These can be estimated from estimations of body movement and ground reaction forces by applying inverse dynamics methods. A major drawback of the existing systems is the restriction to the laboratory environment. Therefore research is required for the development of measurement systems to perform these measurements in an ambulatory environment.

The objective of this study is to give an overview of the possibilities of the forceshoe. The forceshoe is an ambulatory measurement system able to measure the ground reaction force and movement of foot and ankle.

Methods

The forceshoe consists of an orthopaedic sandal equipped with two six-degrees-of-freedom force/moment sensors (ATI-Mini45-SI-580-20, Schunk GmbH & Co. KG) beneath the heel and the forefoot. Moreover, an inertial sensor is rigidly attached to each force/moment sensor (Figure 1).

The estimation of ankle and foot dynamics requires the ground reaction force and movement of foot and ankle to be determined. The ground reaction force is measured by the force/moment sensors beneath the sole of the forceshoe. The movement of foot and ankle is estimated from signals measured by the inertial sensors (Xsens, MTx, Enschede, The Netherlands) connected to the force/moment sensors. A detailed description of the measurement system and methods can be found in [1].

The estimation of center of mass movement is based on fusion of center of pressure data with double integrated ground reaction force data, both estimated from signals measured by the forceshoe. The fusion is based on a frequency domain method, which is described in [2].

Several measurements were performed with the forceshoe. During the measurements, a subject was asked to walk through the gait laboratory while wearing the forceshoe. The accuracy of the ambulatory system was validated by comparing it to a reference system consisting of an optical measurement system and two force plates.

Figure 1. *Picture of the forceshoe with two force/moment sensors beneath the heel and the forefoot and two inertial sensors rigidly attached to the force/moment sensors.*

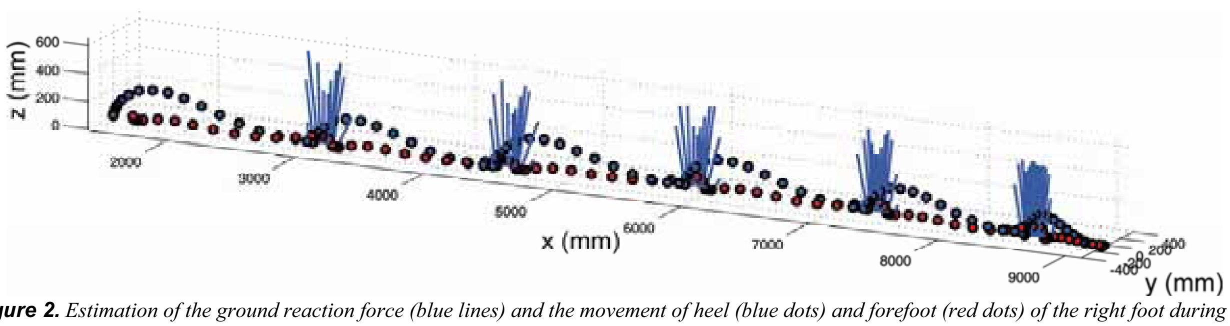


Figure 2. *Estimation of the ground reaction force (blue lines) and the movement of heel (blue dots) and forefoot (red dots) of the right foot during several steps.*

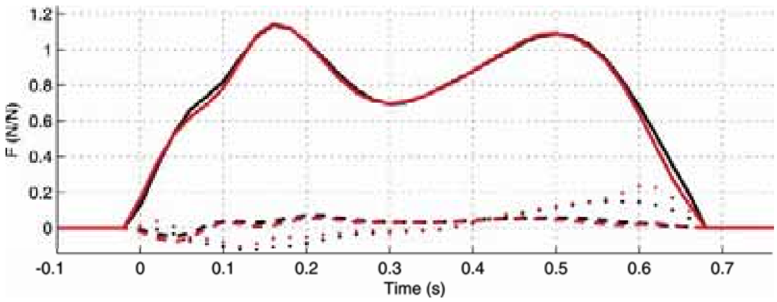


Figure 3. *Three components of the ground reaction force as a function of time estimated by the ambulatory (red) and reference (black) systems (forward: dotted, lateral: dashed, vertical solid).*

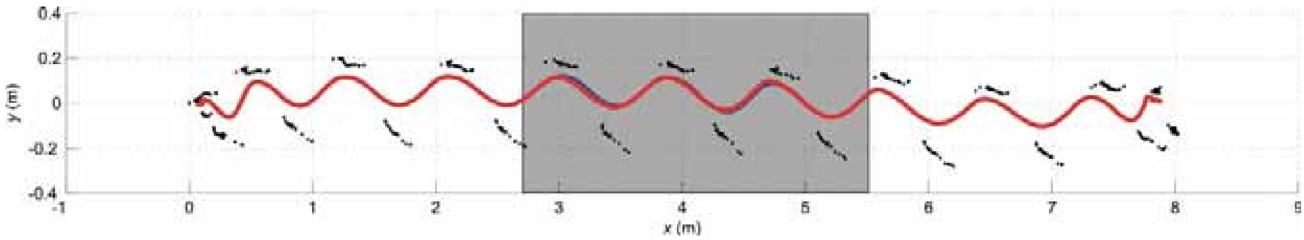


Figure 4. *Estimated center of mass by the ambulatory (red) and reference (blue) systems. The center of pressure is indicated by the black dots on either side of the center of mass, where each dot represents a time sample. Moreover, the measurement volume of the reference system is indicated by the gray area.*

Results

An estimation of the ground reaction force is shown in Figure 3. The signals measured with the forceshoe show good correspondence with the signals measured with the force plate, which is confirmed by the rms difference between the magnitudes of the ground reaction force, being 0.02 N/N or 1.8 % of the maximal magnitude.

Figure 2 shows an integration of the measured ground reaction force with the estimated position of the heel and forefoot sensor. The figure indicates the possibility of the ambulatory measurement system to measure several steps during a single measurement, which is not possible with the reference system. This is also shown in Figure 4, which shows the estimated center of mass movement estimated by the ambulatory (red) and reference (blue) systems. On either side of the center of mass, the center of pressure is indicated by the black dots where each dot represents a time sample. The rms difference between the magnitudes of the center of mass displacement estimated by the ambulatory and the reference measurement systems was 0.025 ± 0.007 m.

Discussion

This study has shown the ability of the forceshoe for ambulatory measurements. Ankle and foot dynamics as well as the movement of the center of mass were estimated and the accuracy was validated using a reference measurement system. A more detailed evaluation with respect to the performance of the forceshoe can be found in [1,2]. Overall, the accuracy of the results obtained with the ambulatory measurement system was comparable to other studies described in literature [3,4].

Acknowledgment

The financial support from the Dutch Ministry of Economic Affairs for the FreeMotion project is gratefully acknowledged.

References

1. Schepers, H.M, Koopman, H.F.J.M., Veltink, P.H., (2007). Ambulatory assessment of ankle and foot dynamics. *IEEE Transactions on Biomedical Engineering* **54**, 895-902.
2. Schepers, H.M. Van Asseldonk, E.H.F., Buurke, J.H., Veltink, P.H. Ambulatory estimation of center of mass displacement during walking. *submitted.*
3. MacWilliams, B.A., Cowley, M., Nicholson, D.E. (2003). Foot kinematics and kinetics during adolescent gait. *Gait and Posture* **17**, 214-224.
4. Eames, M.H.A., Cosgrove, A., Baker, R. (1999). Comparing methods of estimating the total body centre of mass in three-dimensions in normal and pathological gaits. *Human Movement Sciences* **18**, 637-646.

4 years of FreeMotion: towards practical large scale application of ambulatory 3D analysis of human movement

Chris T.M. Baten
(Roessingh Research and Development, Enschede, Netherlands)
FreeMotion Consortium
www.FreeMotion.tk. C.Baten@RRD.nl

Issue

How to transfer the state of the art decision making methods and expertise from elite motion analysis labs in to methods usable for all (peripheral) professionals as to facilitate access to optimal clinical decision making for a much larger group of patients in rehabilitation, ergonomics and sports?

Introduction

Decentralizing health care and a shift from care to prevention prompts for a shift of clinical decision making capacities from specialized motion analysis labs into the hands of the larger community of peripheral - properly educated - professionals in healthcare, ergonomics and sports. Over the last decade ambulatory 3D motion analysis methodology has natured into a potential vehicle for this transfer. This paper presents concepts, methods, technologies, potentials and challenges for ambulatory accurate 3D human motion analysis required for successful transfer resulting from the ongoing FreeMotion research effort executed by a consortium of 10 Dutch research groups and innovative companies [6].

Methods and results

To develop methods for accurate ambulatory 3D analysis of human movement at least the following 4 requirements have to be fulfilled Requirement 1: Portability, large scale applicability and independence of location or external equipment. An ambulatory, portable, wearable, low-cost concept for 3D motion analysis was developed in the form of a complete body worn system for recording 3D human movement, ground reaction forces and muscle activation patterns. This included a wireless infrastructure for peer-to-peer support [1, 3].

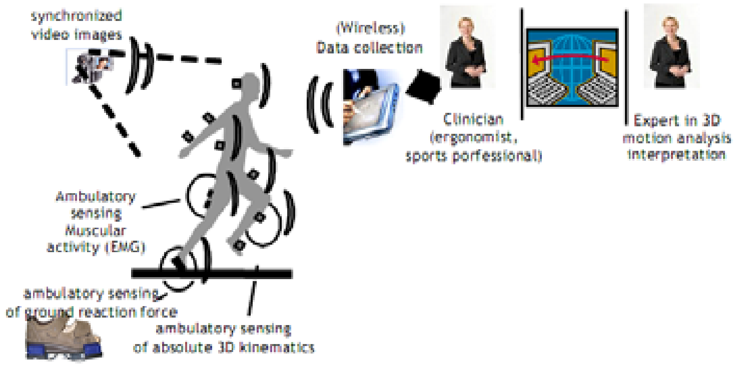


Figure 1. *FreeMotion concept for ambulatory 3D analysis of human movement combining body worn inertial sensors for kinematics, force and EMG, wireless video streams real-time wireless recording and data representation, and modern internet support.*

Figure 2. *Left: Real time viewer beta-GUI example for an application aiming at improving training of specific low back pain patients towards more favorable back use (IC-Coach) Clinical relevant info (ROM, kinematics phase plot) are shown synchronized with context data (video + 3D avatar). Right: Using this real-time application a patient is coached in balancing load exposure of her lower back by a physical therapist in a back school program*

Proceedings of Measuring Behavior 2008 (Maastricht, The Netherlands, August 26-29, 2008)
Eds. A.J. Spink, M.R. Ballintijn, N.D. Bogers, F. Grieco, L.W.S. Loijens, L.P.J.J. Noldus, G. Smit, and P.H. Zimmerman

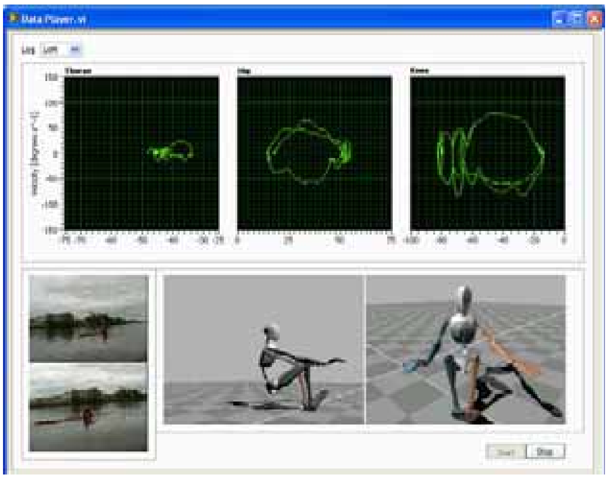


Figure 3. *Example in sports coaching (here varsity rowing), where a sports coach is assisted in real-time monitoring 3D posture, movement and timing of a rower on the water. The coach can view the rower's motion from all sides and study specific timing and kinematics parameters. 2 wireless video streams provide context awareness.*

Requirement 2: Preservation of accurate 3D kinematics assessment. Methods for this was developed (e < 3) based on sensor fusion through optimal (Kalman) estimation of 3D accelerometers, 3D rate gyroscopes and 3D magnetometer signals implemented in mini motion sensor modules for accurate rotational and eventually translational data [5].Requirement 3: Fast, reliable mounting and calibration. For this sensor embedded suits and functional calibration procedures were developed aimed at preventing and eliminating motion artifacts and estimate optimal 3D body segment kinematics from sensor casing kinematics. Great care was taken to minimize patient and clinician burden [2].

Requirement 4: Offer functionality in sensible clinical decision making. A software tool was developed to facilitate clinical sensible application, which records, processes and visualizes in real time 3D motion of up to 10 body segments plus context data (synchronized video of a remotely controlled webcam and 3D full body avatars in 3 simultaneous views). Clinically sensible data representations are added following a tedious interactive analysis of the clinical decision making process [4].

Conclusion

A fully wearable concept of location independent accurate 3D motion analysis in non-elite lab settings seems feasible. From clinical studies evaluating the added value of FreeMotion based tools (e.g. for coaching a-specific low back pain patients into different behavior, for evaluating stroke patients gait patterns, for assessing sports physiotherapy after ACL surgery and for back load monitoring on the job) follows that clinically applicable and useful, robust, affordable, quick applications are feasible Acknowledgement: This paper represents the work of all researchers active in the FreeMotion context. Funding is provided by Dutch Ministry of Economic Affairs and the Innovation Center for Pain Research and the Innovation Center for Rehabilitation Technology, both in Enschede.

Acknowledgements

This paper reflects results of a decade of research by the 10 partners of the FreeMotion research consortium (www.FreeMotion.tk). Sponsoring was provided for parts by the EC, the Dutch Ministry of Economic Affairs, the Dutch Ministry of Health and Education, STW and Senter-Novem, the Netherlands

References

1. Baten CTM, van Moerkerk H, Kingma II et al. (2000),'Ambulatory low back load exposure estimation.' Proceedings of the IEA 2000/HFES 2000.
2. Chris T.M. Baten, M..D. Klein Horsman, W.H.K. de Vries, D.J. Magermans, H.F.J.M. Koopman, F.C.T. van der Helm, P.H. Veltink (2004) 'Estimation of body segment 3D kmiinematics from inertial kinematics', 8 *Int. Congr. On 3D Analysis of Human Movement*, Tampa Fl.
3. Baten, CTM (2006), 'Location free motion analysis', *Proc. of the int. ISEK Conf.*, Torino.
4. Baten CTM, Geerdink F, van Dieen JH (2007),'3D ambulatory motion analysis; already usebla in practical (industrial) physical therapy?', (Dutch laguage), *Special Issue Wetensch. Tijdsch. v. Fysioth*, **117(6)** p233-241.
5. H.J. Luinge, P.H. Veltink, C.T.M. Baten (2007), 'Ambulatory measurement of arm orientation.', *J. Biomechanics* **40**, P 78–85..
6. FreeMotion project: "www.FreeMotion.tk"", 2007.

Analysis of human navigation and manipulation motions

A. Egges
Center for Advanced Gaming and Simulation, Utrecht University, Utrecht, the Netherlands, egges@cs.uu.nl

Introduction

Over the last few decades a lot of research has been done in the field of human motion analysis. A very well-known technique used for motion analysis experiments is the point-light display [3]. Although the point-light display has been a viable tool for such types of experiments, more modern technology, such as a motion capture system, can be used to record human motion in a much more precise manner. There are many advantages to using a motion capture system for motion analysis experiments:

- The recorded data is a precise 3D representation of human joint motions
- No additional video analysis tools are required for retrieving the data
- A motion capture suit with optical markers does not inhibit natural motion
- Recorded motions can be easily visualized on 3D characters

At Utrecht University, we are involved in a variety of motion analysis experiments. One of the most important goals of these experiments is studying combined navigation and manipulation actions, an area of research that is still quite unexplored. Examples of such combined actions are walking to a door and opening it, or walking past a table while picking up an object from it. Our final objective is to define a model that can predict such kinds of combined actions. Recent studies show that very specific motions in the industrial domain can be predicted by a model with a certainty of 90 percent [4]. Although these results are an exception in the study of human motion it also shows that if one have enough information about the subject and the task the subject will perform, one can predict motion. Other remarkable results are shown in recent work by Arechavaleta et al [1,2]. One of their conclusions was that the path humans take to reach a certain goal can be predicted by a clothoid curve, also known as a Curnu spiral, for about 90 percent of the studied takes. Another remarkable result was that the trunk can be described as the 'steering wheel' of human locomotion.

The goal of the research shown in this presentation is to find out whether the results from these previous experiments still hold when applied to combined navigation and manipulation actions. As a test case for this research, we have analyzed the motion of several subjects walking toward a door and opening it (see Figure 1).

Experiment setup

We selected a group of ten subjects (five female, five male) for our experiments. Of the ten there were five left-handed and five right-handed subjects. With each of the subjects, four experiments were done. In the first two experiments, a subject had to move towards the door and open it as if someone rang the doorbell. These experiments were taken from five different start positions. The second experiment differed from the first in the sense that a subject carried a coffee cup in their preferred hand. In the third and fourth experiment the subjects not only walked to the door and opened it, they also passed through it, going in three different directions afterwards. In the first three experiments the door opened towards the subject with the door-latch on the right side. In the last experiment the door-latch was on the left side and the door opened outwards.

Figure 1. *The motion capture lab with the door located in the center.*

Using the results of these experiments, we address five main research questions:

- What is the 'steering wheel' of the human body in case a combined motion action is performed?
- Is the approaching path predictable in a setup of navigation and manipulation?
- When do subjects use the left of right hand to open the door?
- Is the approaching speed influenced by gender, handedness, approaching direction, the direction in which the door opens or the direction in which the subjects leave after walking through the door?
- Is the distance and position relative to the door while opening it influenced by gender, handedness or approaching direction?
- Is the distance to the door while opening it influenced by gender, handedness or approaching direction?

Results

One of the first questions we studied was the steering wheel principle. We investigated whether the trunk is really the steering wheel and whether this still holds in combined navigation/manipulation actions. Our results confirm that the trunk is indeed the steering wheel while navigating. However, this changes as soon as a manipulation action is performed. During manipulation tasks, the upper body of the subject becomes involved in the manipulation action, and thus the trunk can no longer be used as the steering wheel. Additionally, more complicated navigation actions are performed at the same time, such as side-stepping or moving backward. In these cases the relation between the direction of movement and the direction the body parts are facing is lost.

We also analyzed the average distance of the subject to the door while opening it. Our analysis shows that there was no significant relation between the distance to the door and the gender. We have also looked at the role of different body parts used to cover this distance (legs, trunk and arms). Although more thorough analysis is required, the pose seemed to be very different per gender and handedness. Actually, in a few cases the subjects leaned backwards while opening the door.

We are currently in the process of further analyzing the data. During the presentation, a more extensive outline will be given of the results of the analysis and how this translates to the research questions that were formulated in the previous section.

References

1. G. Arechavaleta, J-P. Laumond, H. Hicheur, and A. Berthoz. Optimizing principles underlying the shape of trajectories in goal oriented locomotion for humans. *In IEEE-RAS International Conference on Humanoid Robots.* Genova (Italy), 2006.
2. G. Arechavaleta, J-P. Laumond, H. Hicheur, and A. Berthoz. The nonholonomic nature of human locomotion: a modeling study. Pisa (Italy), 2006. *IEEE / RAS-EMBS International Conference on Biomedical Robotics and Biomechatronics.*
3. L.T. Kozlowski and J.E. Cutting. Recognizing the sex of a walker from a dynamic point-light display. *Perception and Psychophysics*, 21:575–580, 1977.
4. David W. Wagner, Mathew P. Reed, and Don B. Chaffin. Predicting foot positions for manual materials handling tasks. Presented at SAE Digital Human Conference, Iowa City, Iowa, 2005.

Combining manipulation and navigation in virtual environments

B.J.H. van Basten
Center for Advanced Gaming and Simulation, Utrecht University, the Netherlands, basten@cs.uu.nl

Introduction

Character animation plays an important role in games, simulations and movies. Unfortunately, creating animations from scratch is a very tedious job and the resulting quality depends on the skills of the animator. Therefore a lot of research has been done on automatically generating realistic animation to assist the animator.

Although current algorithms are able to generate a wide range of motions, most of them do not take human behavior into account. Consider the following task: picking up a cup from a table while walking passed it. The resulting animations from current algorithms look far from natural because they do not consider factors that influence the path and animation. For instance, the end position of the path might determine which hand will be used for picking up the cup.

In an ideal situation, we would only like to instruct our character *what* to manipulate: our algorithm will generate a realistic path to the object followed by a realistic animation of the manipulation. So, our research is two-fold. We conduct experiments in our motion capture lab to determine a model that yields natural path and postures given specific situations and we develop algorithms that generate animations that are conform this model.

Related work

One of the best-known algorithms for motion synthesis (generation) are the *motion graph* [1] approaches. A motion graph is a directed graph that encapsulates the ways existing motion clips can be re-assembled into new motions. It contains original motion clips (for instance retrieved by motion capture) and generated transitions. These transitions connect two clips of original data by blending from one frame to another frame. In order to automatically determine these transitions several distance comparison metrics can be used to determine to what extent frames resemble each other. Then, after the motion graph is created, a graph walk can be applied to generate a new motion. Many motion graph variations exist that mainly restructure the graph and/or apply better search algorithms to synthesize motions more efficiently.

Unfortunately, the generated motions do not always fit our needs because a target object that we want to pick up might be at other positions than in the concatenated motions. By using motion interpolation [2] techniques we can generate picking up motions that correspond to the target object.

Recently, several methods have been developed that combine motion concatenation (such as the motion graphs) and motion interpolation. Heck *et al.* introduce *parametric motion graphs* [3]. Continuous "spaces" of motions are constructed that are connected in a graph-like structure. These spaces are basically comprised by a set of resembling motions that can be parameterized by interpolation techniques.

Although parametric motion graphs result into a continuous stream of parameterizable motions, it is still not possible to combine motions naturally, such as needed for walking past a table and picking up a cup. To address part of this problem, Heck *et al.* [4] present a technique to attach (splice) a lower-body motion with an upper-body action. Unfortunately, the upper and lower body motions are still fixed and except for some alignment between body parts, human behavior is still not taken into account. Shapiro *et al.* [5] also split the problem into two parts. One part (the locomotion) is fixed; the other part (manipulation) is planned on top of the lower body locomotion. They use path planning techniques to plan the arm movement to manipulate or avoid objects. The locomotion, however, is still fixed and dependencies between the manipulation and navigation are not taken into account.

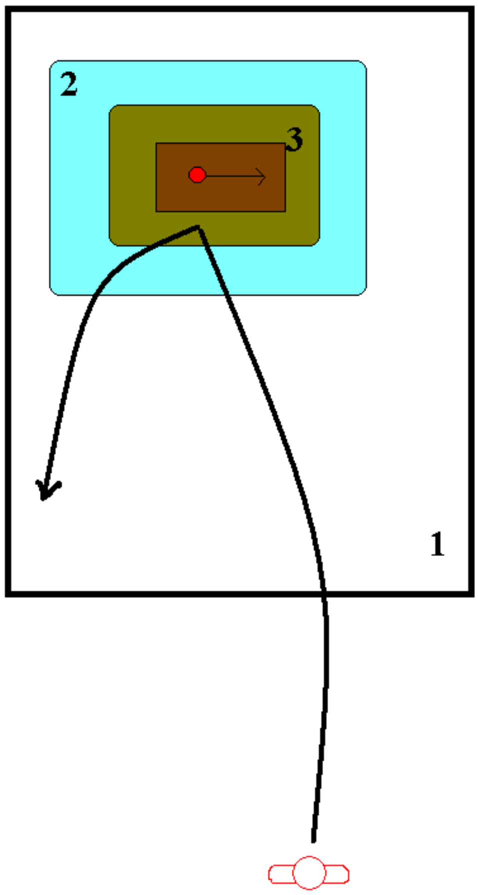


Figure 1. *Several zones can be distinguished in tasks consisting of navigation and manipulation*

Model

When we consider a combination of manipulation and navigation such as depicted in Figure 1 (the character walks to the brown table, moves the red cup to the right and walks back) we can distinguish several zones.

At first, the character walks globally towards the table, when entering zone 1 (*navigation* zone) he adapts his path according to the task and goals he will have in the future. When entering zone 2 (*preparation zone*), the character will prepare his manipulation task, such as already reaching out. In the third zone, the character will perform the actual manipulation.

Many parameters will influence this task. The shapes of these zones and the shape of the path are all dependent on the goal, the starting point of the character, the properties of the table and cup etc. Our research focuses on determining and analyzing the factors that influence the shapes of these zones.

Proceedings of Measuring Behavior 2008 (Maastricht, The Netherlands, August 26-29, 2008)
Eds. A.J. Spink, M.R. Ballintijn, N.D. Bogers, F. Grieco, L.W.S. Loijens, L.P.J.J. Noldus, G. Smit, and P.H. Zimmerman

Once we have a proper understanding of these zones, we analyze the behavior in such a zone itself.

Experiments

In order to determine these zones, we are conducting several human behavior experiments in our motion capture lab. The resulting data does not only provide us with insight in human behavior but also allows us to derive numerical data in order to develop animation algorithms to simulate these kinds of tasks.

References

1. Kovar, L., Gleicher, M., Pighin, F. (2002). Motion graphs. *Proceedings of the 29th Annual Conference on Computer Graphics and interactive Techniques* (San Antonio, Texas, 23-26 July 2002), 473-482.
2. Wiley, D. J., Hahn, J. K. (1997). Interpolation Synthesis of Articulated Figure Motion. *IEEE Comput. Graph. Appl* **17**: 39-45.
3. Heck, R.,Gleicher, M. (2007). Parametric motion graphs. Proceedings of the 2007 Symposium on interactive 3D Graphics and Games (Seattle, Washington, April 30 - May 02, 2007), 129-136.
4. Heck, R., Kovar, L., and Gleicher, M. (2006). Splicing Upper-Body Actions with Locomotion. Computer Graphics Forum **25**: 459-466.
5. Shapiro, A., Kallmann, M., Faloutsos, P. (2007). Interactive motion correction and object manipulation. Proceedings of the 2007 Symposium on interactive 3D Graphics and Games (Seattle, Washington, April 30 - May 02, 2007), 137-144.

Using motion capture data to generate and evaluate motion models for real-time computer animation

H. van Welbergen
Human Media Interaction, University of Twente, Enschede, the Netherlands, welberge@ewi.utwente.nl

Introduction

In the field of computer animation, we are interested in the creation of movement models that make a virtual human (VH) move in a natural way. Ideally, one should not be able to distinguish the movement of such a VH from that of a real human. Furthermore, we want to be able to exert control over such motion in real-time, so that the motion can be adapted or fully generated during interaction with VHs. Obtaining such control in real-time typically comes at the cost of naturalness. That is why VHs that look and move in a very natural way are seen in movies, where all behavior is predefined, while we can not interact with such natural looking and moving humans in real-time.

Our approach to motion generation is bottom up. We start out with motion capture data and replace the motion on a part of the body motion by a movement model. For example, we could replace the lower body movement by a balancing movement model, or we could replace movement of the head and eyes by a gaze model, while the rest of the body is still moved by motion capture. This way, a movement model can be evaluated in isolation in a user test, by comparing it with motion capture data.

Creating motion models

Motion editing techniques use the recorded motion directly in the motion model. The motion is generated using a combination of existing motion recordings. Recordings that are 'close enough' can be concatenated to generate new motion [1], or several motions can be blended to form a desired motion [2]. Control in motion editing techniques is about finding the right motions to combine and finding blend weights that produce a desired motion.

Physical models steer the motion of the body by applying forces in joints. In real-time physical animation, these forces are calculated by models from control theory: a desired state of the body is defined, and the forces are steered so that the body gets closer to this desired state [6]. For example, our physical balance model steers forces in the hips, knees and ankles, in such a way that the body's center of mass moves closer to a desired position in which the body does not fall over. Control in such models is about finding the right control model for a certain task, and about setting parameters in the desired state of the body (such as the height of the hips above the ground and the position of the balance point in our balance model).

Procedural simulation defines mathematical formulas that control motion, given motion time and a set of movement parameters. This can be used to directly control the rotation of joints [3]. A typical application is at a slightly higher level: the movement path of hands through space is defined mathematically to generate gestures [4].

Procedural models and physical models are typically created on the basis of models from biomechanics or behavior science, rather than directly basing them off motion capture. The parameters that steer these models are designed to be intuitive for motion authors, but are often related. Motion capture can serve as a way to find dependencies between these parameters. For example, we have shown that the movement path of the hand decrease linearly with the tempo in a clapping task [5]. A change of one parameter then changes all parameters that are related to it. If more than one parameter is specified, conflicts might arise. These conflicts can be solved in several ways, for example by finding some kind of 'best fit' of parameters values, weighted by their importance.

Evaluating motion models

VHs usually do not have a photo-realistic embodiment. Therefore, if the naturalness of VH animation is evaluated by directly comparing moving humans with a moving VH, the embodiment could bias the judgment. To remedy this, motion captured human movement can be casted onto the same embodiment as the VH. This motion is then compared with generated animation. Typically this is done in an informal way. A motion Turing Test [6] could be used to do this more formally. In such a test, subjects are shown generated movement and similar motion captured movement, displayed on the same VH. Then they are asked to judge whether this was a 'machine' moving or a real human. However, such a human judgment is not sufficient to measure the naturalness of motion. Even if a certain movement is judged as natural, an invalidation of naturalness that is not noticed consciously can still have a social impact [7]. Unnatural moving characters can be evaluated as less interesting, less pleasant, less influential, more agitated and less successful in their delivery. So, while a VH Turing test is a good first measure of naturalness (at least it looked human-like), further evaluation should determine if certain intended aspects of the motion are delivered. Such aspects could include showing emotion, enhancement of the clearness of a spoken message using gesture, showing personality, etc.

We use a movement model that steers a part of the body, and steer the rest of the body using motion capture. We can then compare the motion generated by the movement model combined with motion capture with the same motion generated solely by motion capture in a motion Turing test. In a similar way, we can test if a certain aspect of motion is important for naturalness, by using a model that either removes this aspect, or replaces it by noise. Our method does not only provide us the means to test motion models in isolation, but it also provides meaningful technology to combine motion models with kinematic motion. In a later stage, we plan to use this approach to test combinations of one or motion models that were evaluated to work well in isolation.

Acknowledgements

This research has been supported by the GATE project, funded by the Netherlands Organization for

Scientific Research (NWO) and the Netherlands ICT Research and Innovation Authority (ICT Regie).

References

1. Kovar L., Gleicher M. Pighin F. H. (2002). Motion graphs. *ACM Transactions on Graphics* **21(3)**, 473-482.

2. Wiley D. J., Hahn J. K.(1997). Interpolation Synthesis of Articulated Figure Motion. *IEEE Computer Graphics Applications* **17(6)**, 39-45.
3. Perlin, K. (1995). Real Time Responsive Animation with Personality. *IEEE Transactions on Visualization and Computer Graphics* **1(1)**, 5-15.
4. Chi D. M., Costa M., Zhao L., Badler N. I. (2000). The EMOTE model for effort and shape. *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*, 173–182.
5. van Welbergen H., Ruttkay Z.M. (To appear). On the parameterization of clapping. *Proceedings of the 7th International Workshop on Gesture in Human-Computer Interaction and Simulation (Lisbon, May 26 2007)*
6. Hodgins J. K., Wooten W. L., Brogan D. C., O'Brien J. F. (1995). Animating human athletics. *Proceedings of the 22nd annual conference on Computer graphics and interactive techniques*, 71-78.
7. Reeves B., Nass, C. (1996). The media equation: how people treat computers, television, and new media like real people and places. Cambridge University Press.

Using motion capture to recognize affective states in humans

Nadia Bianchi-Berthouze
UCLIC, University College London, London, UK, n.berthouze@ucl.ac.uk

Humans convey and recognize affective states from a broad spectrum of verbal and nonverbal modalities. An historical focus on face as a primary modality for conveying and recognizing emotions spurred the computer science community to research methods for computer systems to automatically recognize emotions from facial expressions. Today, the Facial Action Coding System (FACS) [1] is the most popular standard for systematically categorising (facial) expressions of emotions. Recent psychology studies, however, have revealed that another form of nonverbal communication, body posture, can prove a very good indicator for various categories of emotion. Whilst these studies have been used rather effectively to enable artificial systems to express affective behaviour through posture (e.g., Sony's AIBO), posture still has no equivalent to FACS and most existing studies use only coarse-grained posture descriptors (e.g. leaning forward, slumping back).

Over the last 5 years, my research has been focused on investigating the extent to which low-level features of body postures provide the information necessary to recognize not only basic emotional states but also more subtle nuances, cross-cultural differences, and affective dimensions. In this paper, I briefly review how I used motion capture to understand how humans recognize affective states from postures and how build automatic affective posture recognition models that could make technology more engaging.

Low-level description of posture

To establish the groundwork for a FACS-like formal model, we used motion capture to record 3D affective gestures from actors of different age, gender and race. Each actor was asked to perform an in-place gesture expressing happiness, sadness, fear or anger. The actors were not forced in their acting, but were allowed to express the emotions in their own natural way. The actors were not allowed to observe each other during performance. Each actor was dressed in a suit with 32 markers on the joints and body segments. 32 markers provide sufficient information to describe the posture. Each gesture was captured by 8 cameras and represented by consecutive frames describing the position of the 32 markers in the 3D space. A total of 109 gestures were collected. For each gesture, we selected the frame (i.e., a static posture) that the actor evaluated as being the most expressive instant of the gesture. Each frame was described using 24 posture features, chosen based on the concept of ''sphere of movement'' used in dance to convey emotion. Direction and volume of the body were described by projecting each marker on the 3 orthogonal planes and measuring the lateral, frontal and vertical extension of the body. Each posture was initially rotated to simulate a frontal view of the posture. The features computed include rotation angles describing head and torso configurations as well as a series of distances between key anatomical landmarks. Each feature was normalized according to the body structure of the actor, i.e., according to the maximal extension of his/her body. For example, the lateral opening of the right arm was computed by the ratio of the distance between the right hand and the left shoulder along the X-direction, and the maximum lateral extension of the arm. In [2], we showed how a trained associative neural network could successfully classify just over 70% of 102 postures extracted from natural human motion capture data. Adding a measure of the direction of the movement to the postural descriptor allowed for a significant improvement (+8%) of postures that showed the lowest inter-observer agreement [3]. In [4], we tested the informational content of the posture descriptors by applying mixed discriminant analysis (MDA) and looking at whether the features could account for different levels (high, low) of three affective dimensions: arousal, valence and action tendency. The results showed a 1% error on arousal, 20% on valence and 25% on action tendency. Finally, we showed that by using both supervised and unsupervised learning mechanisms, nuances of emotions could be recognized from these low-level features, with performance similar to that of human observers [5].

How do people interpret affective posture?

The above motion data enabled us to create faceless computer characters to build an understanding of how people interpret affective postures. Using the above descriptive system, our studies revealed significant effects of both culture [6] and gender [7] on the affective appraisal of body postures. Interestingly, our set of low-level feature descriptors did also provide a mechanistic explanation to recent findings in neuroscience suggesting that the face fusiform area (FFA) – the brain area responsible for facial processing – was involved in processing postural expressions of affect even when facial cues were removed [8]. Indeed, our statistical analysis showed that features related to head configuration (e.g., inclination and rotation of the head) were very important in discriminating between emotions [9] and between nuances of a given emotion in particular [5]. This body of work thus suggests that posture could be used, if not as an alternative to facial expressions, at least in conjunction with facial expressions to provide for finer grain appraisals and increased discriminatory power in the case of ambiguous or incongruent information. This is not the only contribution of posture to our study of emotion in human-machine interaction, however.

Beyond acted postures

Whereas the previous studies relied on acted postures, our current research is concerned with real-world scenarios in which the expressions can be more subtle and mixed. More specifically, the AffectME project (http://www.cs.ucl.ac.uk/staff/n.berthouze/AffectME.html) considers two scenarios: games involving full-body controllers (e.g., Wii), and rehabilitation of patients with chronic pain. The game industry has recently introduced full-body controllers, presumably to help create a more natural and engaging experience for the players. Yet, there have been few studies aimed to understand the relationship between body movement and engagement. In [10], we used using motion capture to quantify differences in movements between players using a traditional game pad controller and those using a full-body controller, and correlated them with measures of engagement [11]. In the clinical study, we are applying the above framework to automatically discriminate between different communicative roles of body movement in chronic pain patients. Although pain, as such, is not an emotion, it is associated with a set of negative emotions that will express at the postural level. Studies in non-verbal behavior and pain have shown that movement in patients convey three different types of information: the physical reaction to pain; the affective experience related to pain; and the search for empathy and attention of solicitous others (e.g., a partner or a practitioner). While we are still at very preliminary stage in this study, our aim is to create a computational model of body movement able to separate such components so as to enable the creation of technology to support patients in self-directed rehabilitation programs. Motion capture systems provide us with a unique source of accurate and rich data to inform the design and the validation of models with use in various areas of human machine interaction.

References

1. Ekman, P. and Friesen, W. (1978). Manual for the facial action coding system. Consulting Psychology Press, Palo Alto.
2. Bianchi-Berthouze, N., Kleinsmith, A. (2003). A categorical approach to affective gesture recognition. *Connection Science* **15(4)**, 259–269.
3. Kleinsmith, A., Fushimi, T., Berthouze, N. (2005). An incremental and interactive affective posture recognition system. *Int. Workshop on Adapting the Interaction Style to Affective Factors*, *Edinburgh, UK*.
4. Kleinsmith, A., de Silva, R., Bianchi-Berthouze, N. (2005). Grounding affective dimensions into posture descriptions. *Int. Conference on Affective Computing and Intelligent Interaction*, 263–270.
5. de Silva, R., Kleinsmith, A., Bianchi-Berthouze, N. (2005). Towards unsupervised detection of affective body nuances. *Int. Conference on Affective Computing and Intelligent Interaction*, 32–39.
6. Kleinsmith, A., De Silva, R., Bianchi-Berthouze, N.: Cross-cultural differences in recognizing affect from body posture. Interacting with Computers, 18 (6), 1371-1389, 2006
7. Kleinsmith, A., Bianchi-Berthouze, N., Berthouze, L. (2006). An effect of gender in the interpretation of affective cues in avatars. *Workshop on Gender and Interaction, Venice, Italy*.
8. de Gelder, B., Snyder, J., Greve, D., Gerard, G., Hadjikhani, N. (2003). Fear fosters fight: A mechanism for fear contagion when perceiving emotion expressed by a whole body. *PNAS* **101(47)**, 16701–16706.
9. de Silva, R., Bianchi-Berthouze, N. (2004). Modeling human affective postures: An information theoretic characterization of posture features. *J. Computational Agents and Virtual Worlds* **15**, 269–276.
10. Bianchi-Berthouze, N., Whan Woong, K., Patel, D. (2007). Does body movement engage you more in digital play? And why? *International Conference on Affective Computing and Intelligence Interaction, Lisboa (Portugal)*, LNCS **4738**, 102-113.
11. Cairns, P. Cox, A., Berthouze, N., Dhoparee, S. (2006). Quantifying the experience of immersion in games. *Workshop on Cognitive Science of Games and Gameplay, Vancouver, Canada*.

Online Segmentation of Continuous Gesturing in Interfaces

F.W. Fikkert, P.E. van der Vet, and A. Nijholt
Human Media Interaction, Department of Computer Science,
University of Twente, Enschede, The Netherlands,
{f.w.fikkert|p.e.vandervet|a.nijholt}@ewi.utwente.nl

Segmentation of Continuous Gesturing

An increasing demand exists for more intuitive ways to interact with ever larger displays. Natural interfaces fulfill this demand by directly analyzing, reacting to and reasoning about observed human behavior. We focus on gesturing which is a key modality in a more complex natural interface. Most gesture recognition research today focuses on computer vision techniques. Although promising, such techniques are not yet mature enough for markerless, robust and in-the-field deployment. Instead, we deploy a more down-to-earth solution by using existing motion capture systems. Our goal is to develop methods for automatic online segmentation and interpretation of continuous gesturing. We propose a two-stage approach. First, we use commodity hardware equipped with accelerometers and buttons that explicitly mark gesture boundaries. Experience gained there feeds our second stage in which we investigate motion trajectories and their meaning. In the second stage we use a combination of a full body motion capture suit and two data gloves.

Motion segmentation – extracting distinct units from continuous motion – is hard because the boundaries are both subjective and highly sequence-dependent. More so, even predefined gestures are made with variation. Individual users are not fully consistent and two users do not make exactly the same gestures [1]. Applications of automatic gesture segmentation deal with this variety by basing start and end cues on local minima or turn points in the trajectory [2]. This approach is applied often, with varying success, for more coarse, whole body movements such as dancing [3] and conducting music [4]. For finer motion such as sign language [5], extensive machine classifier solutions are mostly used. They are fed with the whole motion sequence. These solutions require extensive training and their computational load prevents real-time motion analysis.

Strong, familiar metaphors are the key of a natural interface [6]. We have devised a series of user experiments in which we hope to discover a gesture repertoire consisting of such metaphors [7]. We distinguish two semantically separate classes of metaphors in gesture-based interfaces. The first class involves manipulations of virtual 3D objects as if they were tangible. This form of interaction is directly based on the way humans interact with their everyday surroundings. Examples are picking up, rotating and zooming of a 3D mesh. The second class involves more abstract metaphors that are ideally inspired by everyday activities [8]. Gesturing to throw an item away means to delete it, for example. This class contains the six communicative gesture classes defined by [9]. Clearly, these two classes are strongly linked. Consider a case in which a person selects two visualizations – by picking them up – and then analytically relates them to each other – by moving them together. The key is to both find these metaphors and to recognize the intended meaning of the gestures.

Approach and results

To discover gesture boundaries and gesture meaning we propose a two-stage approach. In the first stage, we have collected trajectories and explicit gesture boundaries using commodity hardware: the WiiRemote controller from Nintendo's Wii game console. Gesture trajectories and boundaries were gathered through a user study. Users interacted with both a 3D mesh and a collection of 2D images. Trajectories were explicitly bounded using the buttons on the WiiRemote. We will show that these explicitly marked gesture boundaries will feed our segmentation method. This is shown by comparing multiple machine classifier techniques that are trained on these data to recognize the gesture boundaries. For our research, the WiiRemote functions as a stepping stone for high-resolution full body motion capture that we will deploy in stage two. We work towards this complexity because human movement is an effort of the whole body. This system will combine a full body motion tracking suit with a pair of CyberGloveII data gloves.

Acknowledgements

We are grateful for the continued input in our research from the Microarray Department of the University of Amsterdam. This work is part of the BioRange program carried out by the Netherlands Bioinformatics Centre (NBIC), which is supported by a BSIK grant through the Netherlands Genomics Initiative (NGI).

References

1. J. Epps, S. Lichman, and M. Wu. (2006) A study of hand shape use in tabletop gesture interaction, presented at CHI '06: *CHI '06 extended abstracts on Human factors in computing systems*, Montreal, Quebec, Canada
2. J. Barbič, A. Safonova, J.-Y. Pan, C. Faloutsos, J. Hodgins, and N. Pollard. (2004) Segmenting motion capture data into distinct behaviors. presented at GI '04: *Proceedings of Graphics Interface 2004*, London, Ontario, Canada
3. K. Kahol, P. Tripathi, S. Panchanathan, and T. Rikakis. (2003) Gesture segmentation in complex motion sequences. presented at *ICIP 2003 - International Conference on Image Processing.*
4. T.-S. Wang, H.-Y. Shum, Y.-Q. Xu, and N.-N. Zheng. (2001) Unsupervised Analysis of Human Gestures. in Advances in Multimedia Information Processing — PCM 2001, vol. 2195/2001, *Lecture Notes in Computer Science: Springer Berlin / Heidelberg* pp. 174-181.
5. S. Ong and S. Ranganath, Automatic Sign Language Analysis: A Survey and the Future beyond Lexical Meaning, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **27**, pp. 873-891, 2005.
6. T. Ni, G. Schmidt, O. Staadt, M. Livingston, R. Ball, and R. May. (2006) A Survey of Large High-Resolution Display Technologies, Techniques, and Applications, presented at *Proceedings of IEEE Virtual Reality 2006*, Alexandria, VA
7. W. Fikkert and P. van der Vet. (2007) Towards a Gesture Repertoire for Cooperative Interaction with Large Displays, presented at *7th International Workshop on Gesture in Human-Computer Interaction and Simulation* (GW 2007), Lisbon, Portugal
8. T. Grossman, R. Balakrishnan, G. Kurtenbach, G. Fitzmaurice, A. Khan, and B. Buxton. (2002) Creating principal 3D curves with digital tape drawing, presented at CHI '02: Proceedings of the SIGCHI conference on Human factors in computing systems, Minneapolis, Minnesota, USA
9. A. Kendon. (2004) Gesture: Visible Action as Utterance

The influence of gender stereotype priming on social action

E. Ngubia Kuria[1], Luisa Sartori[2], Castiello Umberto[2], and Raffaella I. Rumiati[1]
[1]*Cognitive Neuroscience Sector, SISSA, Trieste, Italy*
[2]*Department of General Psychology, University of Padova, Padova, Italy*

Abstract

The study investigated whether kinematics during social actions can be influenced by the gender of the actor or of the partner. Participants were requested to reach towards, grasp and place the stimulus either in a concave container (Non-social condition) or in the hand of a partner which could be either a female or male (Social condition). The key result is that gender modified the kinematics of the reach-to- grasp movement. In particular, participants displayed faster movement and specific kinematics when interacting with partners of the opposite sex. We contend that these results reflect how gender stereotypes affect the motor aspects of social interaction.

Methods

Participants

Fifteen students, 8 women (mean age = 24.7 years, SD = 3.8), took part in the experiment. All participants were right-handed, had normal vision and were naïve as to the purpose of the experiment.

Stimulus

The stimulus was an egg-shaped object located in front of the participant at a distance of 25 cm from the hand starting position along the midsagittal plane.

Apparatus

The working surface was a rectangular table (150 x 100 cm), and the subject was seated on a height-adjustable chair. Before each trial, participant put their right hand on a starting pad (brown velvet cloth 7 x 6 cm). The pad was attached 3 cm away from the edge of the table on the mid-sagittal axis 15 cm anterior to the subject's midline. Another pad was placed 28 cm at the far right side of the subject, on which either the experimenter's hand, or a small concave container rested. Infrared reflective markers (0.25 mm diameter) were taped to the following points on the participants' right upper limb: (1) wrist – dorsodistal aspect of the radial styloid process; (2) thumb – ulnar side of the nail; and (3) index finger – radial side of the nail. Markers were fastened using double-sided tape. Movements were recorded using an ELITE motion analysis system (Bioengineering Technology & Systems [B|T|S]). Four infrared cameras (sampling rate 100 Hz) placed 120 cm away from each of the four corners of the table captured the movement of the markers in 3D space. Co-ordinates of the markers were reconstructed with an accuracy of 0.2 mm over the field of view. The standard deviation of the reconstruction error was 0.2 mm for the vertical (Y) axis and 0.3 mm for the two horizontal (X and Z) axes.

Procedure

The task consisted of lifting the stimulus, and placing it to a designated position in three experimental conditions:

1. *Non-Social Condition.* Each participant was requested to reach towards the stimulus, grasp it and place it on a small round platform (12 cm diameter) positioned on the end-pad located at his/her right side.

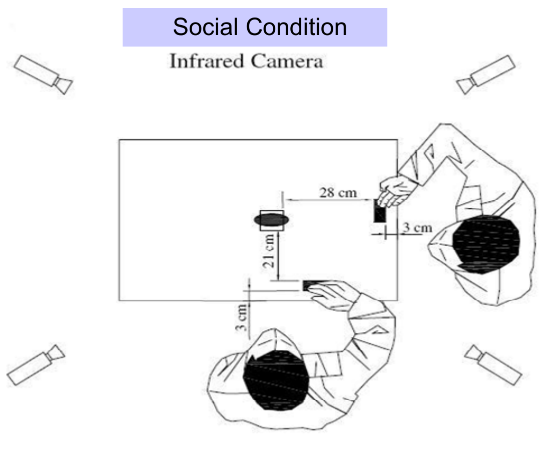
2. *Social Condition – same gender.* Each participant was requested to reach towards the stimulus, grasp it and place it on the co-experimenter's hand. The experimenter was of the same gender as the participant, either female or male.

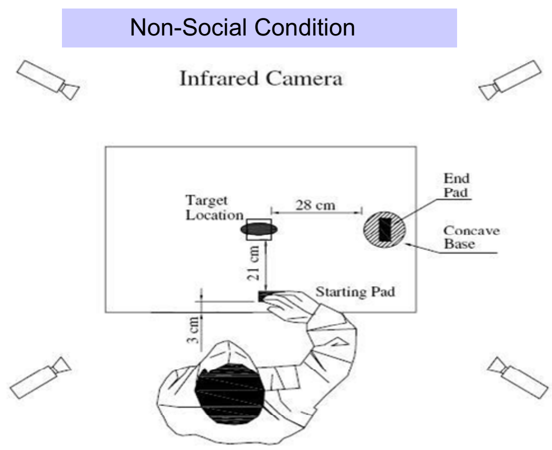
3. *Social Condition – different gender.* Each participant was requested to reach towards the stimulus, grasp it and place it on the co-experimenter's hand. The experimenter was of a different gender as the participant.

For all conditions participants were requested to start the action after a tone (880 Hz/200 ms) was presented. The order of conditions and trials was randomized across participants. For each condition, participants performed 30 trials (90 trials in total).

Data Processing

The ELIGRASP software package (B|T|S|) was used to analyse the data and provide a 3-D reconstruction of the marker positions as a function of time. The data was then filtered using a finite impulse response linear filter (transition band = 1 Hz, sharpening variable = 2, cut-off frequency = 10 HzD'Amico and Ferrigno, 1992). Following this operation, for the reaching component the spatial trajectory and tangential speed of the wrist marker were computed. For the grasp component the distance between the markers located on the index finger and the thumb and the velocity of finger aperture were computed. Tangential speed data were used to determine the onset and offset of the movement using a standard algorithm (threshold for movement onset and offset was ~ 5 cm/s). Temporal data was normalized (as a percentage of movement duration) in order to avoid possible differences due





to possible differences in movement speed between females and males participants.

Data analysis

The movement sequence of each participant was segmented in two action steps: (a) reach towards and grasp the stimulus, (b) lifting the stimulus and transporting it to the required position. Data analysis focused on the first movement phase. This is because this phase was common in all experimental conditions, and possible differences (if any) in social intention should be already evident during this movement phase as previously demonstrated (Becchio *et al.*, 2007a; 2007b). The means of kinematic parameters were entered into a mixed analysis of variance (ANOVA) with experimental condition (non-social, social-same gender, social-different gender) as a within-subjects factor and participants' gender (male, female) as a between-subjects factor. Post-hoc comparisons were carried out using simple effects (Bonferroni corrected, alpha level = $p < 0.05$). Data were preliminarily checked for normality, sphericity (Mauchly test), univariate and multivariate outliers with no serious violations noted.

Results

Interaction between experimental condition and participants' gender was significant for the amplitude of maximum peak wrist velocity [$F(2,26) = 3.39$, $p < 0.05$] and for the time at which the maximum velocity of finger opening was reached in both absolute [$F(2,26) = 5.93$, $p < 0.05$] and relative [$F(2,26) = 3.1$, $p < 0.05$] terms. As shown in Figure 1a, when the participants' action towards the target object implied a social interaction with a partner of a different gender, movement speed was faster ($p_s < 0.05$) and the time at which the fingers opened was anticipated ($p_s < 0.05$; Fig. 1b). Of interest, for the female participants the pattern of anticipation at the level of finger opening did not differ when they interacted with both male and female partners (Fig. 1b) whereas males' differed when they related with female partners.

References

1. Becchio C., Sartori L., Bulgheroni M., Umberto C. (2007). The case of Dr. Jekyll and Mr. Hyde: A kinematic study on social intention. *Consciousness and Cognition, April 17, 2007*
2. Choi J, Amstrong T.J, (2006). Quantitative analysis of finger movements during reaching and grasping tasks. *16th IEA Meeting*
3. Corston, R., Colman A.M., (1996). Gender and social faciliatation effects on computer competence and attitudes towards computers. *Journal of Educational and Computing Research* **14***,* 171-183
4. D'Amico M, Ferrigno G., (1992). Comparison between the more recent techniques for smoothing and derivative assessment in biomechanics. *Medical and Biological engineering and computing* **30**,193-204
5. Georgiou, J., Becchio, C., Glover, S., & Castello, U., (2007). Different action patterns for cooperative and competitive behaviour. *Cognition,* 415–433
6. Gneezy, U., Niederle, M., Rustichini, A., (2003). Performance in competitive environments: Gender Differences. *Quarterly Journal of Economics,* 1049-1074
7. Gneezy, U., Rustichini, A., (2004). Gender and Competition at a Young Age. *The American Economic Review, 2, 377-381*
8. Inzlicht, M., & Ben-Zeev, T., (2000). A threatening intellectual environment: Why females are susceptible to experiencing problem solving deficits in the presence of males. *Psychological Science* **11***,* 365–371
9. Murphy M.C, Steele C.M, Gross J.J., (2007) A Signalling Threat; How Situational Cues Affect Women in Math, Science, and Engineering Settings. *Psychological Science* **18**,879–889
10. Rand M.K, Linda M.S, Stelmach G.E. (2006). Effect of speed manipulation on the control of aperture closure during reach-to-grasp movements. *Experimental Brain Research* **174**, 74-85
11. Van Vugt M, David De Cremer, Dirk P. Janssen (2007). Gender Differences in Cooperation and Competition: The Male-Warrior Hypothesis. *Psychological Science* **18**, 19-23
12. Wallace S.A, Weeks D.L (1988). Temporal constraints in the control of prehensile movement, *Journal of motor Behaviour,* 81-105
13. Wing A.M, Turton A, Fraser C. (1986). Grasp size and accuracy of approach in reaching. *Journal of motor Behaviour,* 245-260
14. Zajonc, R.B. (1965). Social Facilitation. *Science* **16** (149), 269-274

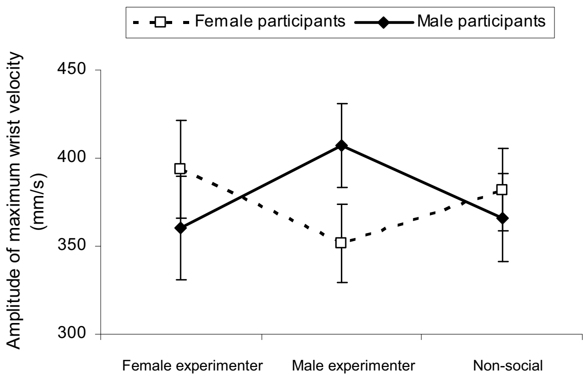


Figure 1a

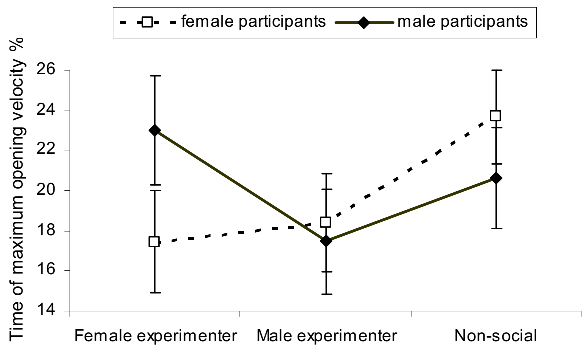


Figure 1b

Measuring recovery after brain and spinal injury in rodents and non-human primates Symposium

Michael S. Beattie and Jacqueline C. Bresnahan
University of California, San Francisco, USA

Preclinical evaluation of treatment strategies aimed at improving outcomes after CNS injury require reliable outcome measures of neurological recovery that are relevant to human clinical trials. This symposium will examine a number of behavioral measures used to evaluate neurological recovery in animals after CNS injury, including the "BBB" locomotor scale and autonomic functional outcomes for spinal cord injury (SCI), tests for forelimb function in brain and spinal injury models in rodents, as well as forelimb tests for non-human primates, and the use of a swimming test for evaluation of recovery after rat SCI. In the final presentation, mathematical methods for evaluating the usefulness of all these measures, and their relationship to underlying mechanisms of repair and recovery will be discussed, and the development of a large database and data mining approach to improving preclinical neurological resting for CBS injury will be presented.

(The symposium organizers' laboratory is supported by the US National Institutes of Health, the Roman Reed Fund of the State of California, The C.H. Neilsen Foundation, and the New York State Spinal Cord Research Board- Center of Research Excellence.)

Symposium contents

Measuring a broad spectrum of clinically relevant outcomes after experimental spinal cord injury

Jacqueline C. Bresnahan and Yvette S. Nout
Brain and Spinal Injury Center, Dept. of Neurological Surgery, University of California, San Francisco, USA
Jacqueline.bresnahan@ucsf.edu

Measuring recovery of function after central nervous system injury is an important aspect of experimental models of brain and spinal cord injury. Our laboratory has been focused on developing models, especially for spinal cord injury, for many years. The purpose of the models is to mimic the neurological condition in humans and to provide a test ground for evaluating treatments which then could be brought to clinical trials. For spinal cord injury, the consequences for humans include loss/reduced motor function with altered reflexes and spasticity, sensory loss and aberrant sensations (e.g. pain syndromes), and autonomic dysfunctions including autonomic dysreflexia, and altered eliminative and sexual functions. These functional changes significantly affect a person's quality of life and ability to work, therefore the full range of these deficits should be modeled, and treatments focused on the multiple endpoints they represent. In a recent survey by Anderson [1], people with SCI rated bladder and bowel function most highly as a critical target for treatment; those with paraplegia also rated recovery of sexual function as a very high priority, whereas those with quadriplegia rated return of hand function as their highest priority. The experimental spinal cord injury literature has only recently begun to recognize these functions as important targets and most studies still evaluate locomotor function as the primary outcome measure, as this is the most easily tested and quantified.

There are many ways to measure locomotor function. The simplest is an open field evaluation using an observational rating scale (BBB) that covers the gamut of possible locomotor functions from total paralysis (no movement) to normal function. The usefulness of this measure is that it provides a general categorization of functional ability. The test has good inter-rater and inter-laboratory reliability, and it can be used to determine when other tests which require a specific level of recovery may be applied. For example, if animals are able to step, runway tests with solid floors (e.g. the CatWalk [2]), or horizontal ladder tests or grid walking tests can be used to measure more refined locomotor abilities. Other tests with a broad range of performance outcomes include swimming tests (e.g. the Schnell swimming test and the Louisville swim score), and the inclined plane test which measures the ability of an animal to maintain position on a platform oriented at increasingly steep angles. Locomotor ability on a treadmill moving at different speeds is also frequently used for kinematic analysis of recovery of hindlimb function, and correlated with EMG patterns. Such analyses frequently show dysregulation of muscle activity which alters the kinematic pattern, even in humans after SCI (e.g. Maegele et al., [3]).

Following thoracic contusion spinal cord injury in rats, in addition to locomotor dysfunction, significant disruption of bladder, bowel and sexual function occurs and can be measured. Hyperactivity/spasticity in external urethral and anal sphincter muscle activity leads to a dyssynergia with autonomically innervated smooth muscle activity, and a concomitant disruption of bladder and bowel function (e.g. [4,5]). Interestingly, these and other functions exhibit recovery profiles over time courses comparable to the more commonly tested locomotor function [4]. However, measuring such functions in freely moving animals is difficult at best. Recently, we have developed methods for measuring autonomic outcomes after injury in awake freely moving animals using implantation of telemetric pressure transducer devices in collaboration with Dr. Markus Schmidt [6,7]. If placed in the corpus spongiosum of the penis, patterns of pressure changes can be correlated with not only penile reflex activity, but micturition as well. Pressure measurements in a variety of behavioral contexts (ex copula reflex testing, non-contact erection test, copulatory test) and video-monitoring of micturition over 24 hr testing periods were recorded and matched to observed erectile and micturition events. There was excellent correspondence between the pattern of pressure recordings and observed erectile and micturition events suggesting that telemetric recordings of CSP pressure can provide quantitative assessments of these events in freely moving animals. In addition, the pattern was affected by spinal cord injury, initially showing a significant reduction in activity which recovered over time, paralleling the recovery of the behaviors. However, the patterns recorded indicated development of hyperactivity in the bulbospongiosus muscle similar to that seen in the external urethral and anal sphincter muscles observed after spinal cord injury. These studies as well as those to be described subsequently in this symposium, show the feasibility of measuring a broad variety of functional outcomes after experimental SCI and suggest alternative analytic approaches to measuring improved recovery after spinal cord damage.

Supported by grants from the NIH (NS31193) and NYState CoRE.

References

1. Anderson, K.D. (2004) Targeting recovery: Priorities of the spinal cord injured population. *Journal of Neurotrauma* **21**, 1371-1383.
2. Hamers, F.P.T., Koopmans, C.G. and Joosten, E.A.J. (2006) CatWalk-assisted gait analysis in the assessment of spinal cord injury. *Journal of Neurotrauma* **23**, 537-548.
3. Maegele, M., Muller, S., Edgerton, V.R., Harkema, S.J. (2002) Recruitment of spinal motor pools during voluntary movement versus stepping after human spinal cord injury. *Journal of Neurotrauma* **19**, 1217-1229.
4. Holmes, G.M., Van Meter, M.J., Beattie, M.S., Bresnahan, J.C. (2005) Serotonergic fiber sprouting to external anal sphincter motoneurons after spinal cord contusion. *Exerimental. Neurology*, **193**, 29-42.
5. Pikov, V., Wrathall, J. (2001) Coordination of the bladder detrusor and the external urethral sphincter in a rat model of spinal cord injury: effect of injury severity. *Journal of Neuroscience* **21**, 559-569.
6. Nout, Y.S., Bresnahan, J.C., Culp, E., Tovar, A.C., Beattie, M.S., Schmidt, M.H. (2006) Novel technique for monitoring micturition and sexual function in male rats using telemetry. *American Journal of Physiology; Regulatory, Integrative, and Comparative Physiology* **292**, 1359-1367.
7. Nout, Y.S., Schmidt, M.H., Tovar, C.A., Culp, E., Beattie, M.S., Bresnahan, J.C. (2005) Telemetric monitoring of corpus spongiosum penis pressure in conscious rats for assessment of micturition and sexual function following spinal cord contusion injury. *Journal of Neurotrauma* **22**, 429-441.

The Schnell Swim Test (SST) to measure motor function and recovery in spinal cord injured rats

Mirjam Gullo, Eva Hochreutener, Dina Schnell, Jeannette Scholl, Martin E. Schwab, and Lisa Schnell
Brain Research Institute, University of Zurich and Department of Biology, Swiss Federal Institute of Technology Zurich, 8057 Zurich, Switzerland, gullo@hifo.uzh.ch

Impaired functions in spinal cord lesioned rodents can be measured by a variety of different behavioral tests which range from elementary locomotor analysis to more demanding assays to analyze motor as well as sensory functions [1, 2, 3, 4, 5, 6]. Some of these tests are based on observation and rating analyses (BBB score, narrow beam), other tests require videotaping and elaborate quantifications (horizontal ladder rung walk, CatWalk®). Unfortunately, none of these tests can be strictly associated with deficits/recovery of a specific fiber system and combinations of different tests are needed to obtain reliable and unbiased motor function values.

For some tasks to be performed properly by the test animal, substantial sensory feedback such as load bearing is necessary (ladder rung etc). When using walking as a functional readout, the disadvantage is that injured animals have to perform rhythmic leg movements while having to overcome gravity. Full weight bearing has been shown to limit recovery of walking in animal models and also in patients where gait training on a treadmill with assistance in weight support is commonly used [7, 8, 9].

In this behavioral test we took advantage of the swimming motion of adult rats. Swimming is a natural behaviour of rats and the buoyancy provided by water enables them to perform locomotor movements without having to support their body weight. Since laboratory animals are normally confined to their cages, the interference with the test by self-training is not an issue. After a short training period, the animals swim without any apparent aversion and the test is hardly influenced by symptoms of stress. Motivation is provided in our setting as the animals prefer to swim straight for the platform. Swimming rats move only their hindlimbs while being supported by the buoyancy of water. Rats with different spinal cord lesions causing large or small deficits in hindlimb and tail functions, show compensatory strokes by the forelimbs. For the development of a reproducible test, velocity, forelimb strokes, hindlimb performance and tail movement were evaluated.

Velocity

The time needed to swim the marked distance of 60 cm was measured by counting the number of frames. The mean value of three runs obtained from trained animals one day before surgical intervention was taken as the baseline (32 ± frames, equal to 1.33 sec per 60 cm). This reading was assigned 100 ± %, equaling normal performance.

Forelimb strokes

The number of forelimb strokes per run was counted using video slow motion settings. Unlesioned, trained animals rarely use their forelimbs and a maximum of 2 strokes per 60 cm swimming distance can be considered the normal value. This represents a 100% performance.

Based on our results, velocity and forelimb strokes, two objectively measurable parameters, correlate already with the extent of the anatomical lesions and allow the assessment of recovery/rehabilitation after treatment. However, additional analyses can be added.

Tail Movement Score

During swimming the rat tail moves in a sinusoidal pattern with strong and regular movements. The impairment of this movement was evaluated and assigned a score from 1 to 4. A score of 4 represents normal tail movements in unlesioned animals, while a completely paretic tail with very rare twitching was given a score of 1 point, and a total absence of movement a score of 0.

Hindpaw Analysis

The normal position of the hindpaws of a swimming rat is close to the body axis. After spinal cord injury, this positioning deviates to a more lateral position and was scored accordingly. The normal swimming position of an intact rat was assigned a score of 4, while the most lateral position was given a score of 1 point. No movement of the hindlimbs (complete paresis) equals a score of 0.

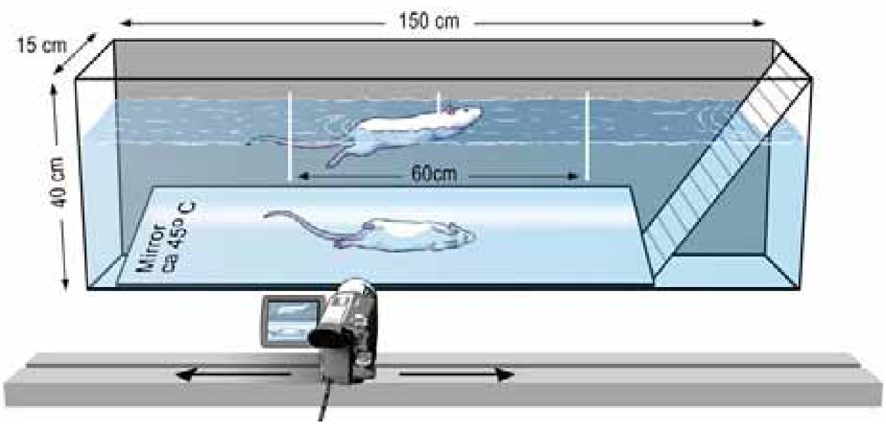


Figure 1. *Swim basin and camera*

Hindlimb-Forelimb Coordination

To judge coordination between hind- and forelimb movement, scoring criteria were used to develop a range from 5 (normal swimming pattern) to 0 (no movement of forelimbs, no movement of hindlimbs).

Inter-hindlimb coordination

Inter-hindlimb coordination was measured in frame-by-frame mode. The time stretch between the backward extension of the left and the backward extension of the right hindlimb was measured by noting the number of frames elapsed during each of 10 cycles. The differences between the numbers obtained between these 10 hindlimb strokes (5 for each leg = 9 intervals) were added and the mean of three runs converted to a percentage of the normal performance.

Summary

Swimming as a behavioral test can be used to obtain objective results in the analysis of motor function after various types of spinal cord lesions. For mild to moderate bilateral dorsal lesions, the assessment of velocity and forelimb strokes can be sufficient to obtain reliable results which correlate with the extent of the lesion or with the degree of recovery after treatment. Swimming also allowed the novel observation in rats of muscle spasms in spinal cord injured rats during the course of recovery (Gonzenbach, in prep.). Such spasms, also seen in paraplegic patients, can be measured in the in swimming rat by determining their body angle or even more accurately by electromyography (EMG). The evaluation of kinematics during swimming is currently under investigation.

References

1. Basso DM, Beattie MS, Bresnahan JC. (1995) A sensitive and reliable locomotor rating scale for open field testing in rats. *J Neurotrauma.* **12(1)**:1-21
2. Metz GA, Merkler D, Dietz V, Schwab ME, Fouad K. (2000) Efficient testing of motor function in spinal cord injured rats. *Brain Res*.**883(2),**165-77.
3. Basso DM, Fisher LC, Anderson AJ, Jakeman LB, McTigue DM, Popovich PG.(2006)Basso Mouse Scale for locomotion detects differences in recovery after spinal cord injury in five common mouse strains.*J Neurotrauma*.**23(5)**:635-59.
4. Muir GD, Webb AA. (2000) Mini-review: assessment of behavioural recovery following spinal cord injury in rats.*Eur J Neurosci.* **12(9):**3079-86
5. Hamers FP, Lankhorst AJ, van Laar TJ, Veldhuis WB, Gispen WH. (2001) Automated quantitative gait analysis during overground locomotion in the rat: its application to spinal cord contusion and transection injuries. *J Neurotrauma.* **18(2):**187-201.
6. Kunkel-Bagden E, Dai HN, Bregman BS. (1993) Methods to assess the development and recovery of locomotor function after spinal cord injury in rats. *Exp Neurol*. **119(2)**,153-64.
7. Cha J, Heng C, Reinkensmeyer DJ, Roy RR, Edgerton VR, De Leon RD. (2007) Locomotor ability in spinal rats is dependent on the amount of activity imposed on the hindlimbs during treadmill training. *J Neurotrauma.* **24(6),**1000-12.
8. Aoyagi D, Ichinose WE, Harkema SJ, Reinkensmeyer DJ, Bobrow JE. (2007) A robot and control algorithm that can synchronously assist in naturalistic motion during body-weight-supported gait training following neurologic injury. *IEEE Trans Neural Syst Rehabil Eng.* **15(3)**:387-400.
9. Dietz V, Duysens J. (2000) Significance of load receptor input during locomotion: a review. *Gait Posture.***11(2)**,102-10.

Skilled Limb Use in Rat Models of Human Neurological Disease

Gerlinde A. Metz
Canadian Centre for Behavioural Neuroscience, University of Lethbridge, Lethbridge, Alberta Canada T1K 3M4, gerlinde.metz@uleth.ca

Introduction

Skilled movements are voluntary movements requiring irregular motor patterns. Skilled movements are characterized by a complex sequence of movement components, which can be observed in rats eating specialty food items, such as sunflower seeds or crickets, or when navigating across difficult territory [5]. Tests of skilled movement are of increasing importance as sensitive assessment tools for rat models in pre-clinical studies of neurological disease.

A number of formal tests for skilled movements have been developed in order to allow objective and reproducible evaluation of limb function. Tests of skilled movement have the advantage to produce a variety of different parameters, such as end point measures and qualitative measures. The majority of skilled movement tests focus on the analysis of forelimb use, such as grasping and retrieving food items, or applying force to retrieve an object. More recently, a task for skilled walking has been developed, which allows assessment of hind limb functions as well as forelimb functions.

Skilled Reaching as a Tool to Assess Recovery and Compensation

One of the most sensitive tests for motor control and postural adjustments in rat models of neurological disease is the single pellet reaching task. The task is designed in a way to allow for measuring reaching success while at the same time performance can be filmed and scored frame-by-frame [1,4]. In this task, rats are trained to extend their preferred forelimb through an opening to grasp and retrieve small food pellets. Once the rat has obtained the food pellet, it withdraws the paw through the opening and sits back on its haunches to consume the food.

The sequence of components comprising the rat's skilled reaching movement is relatively fixed [1]. Thus, rats possess limited ability to modify movement components to adapt to a changing context. Consequently, even a subtle brain damage will permanently compromise reaching movement performance. Reaching movements are differentially affected by discrete lesions of motor cortex and corticospinal tract, basal ganglia, dorsal columns and red nucleus. This makes skilled reaching a useful tool for the study of disease conditions such as spinal cord injury, stroke, and Parkinson's disease. In addition, physiological conditions such as stress, strain, sex and aging affect both reaching success and qualitative movement performance.

Recent studies have shown that skilled reaching movements are useful tools for distinguishing between genuine functional recovery and compensation. Through practice animals with brain damage might show considerable improvement in reaching success [4], however, qualitative analysis of reaching movements might still reveal permanent deficits. This indicates that rats are able to develop successful alternative movement strategies to overcome primary motor deficits. Thus, descriptive movement analysis represents an important supplementary technique to determine compensatory behavior and permanent motor deficits.

The relevance of skilled reaching tasks for rodent models of neurological disease has been established. The study of skilled reaching in rats is especially useful because reaching movements in rats and humans are homologous [6]. Comparisons have shown that the essential components of skilled reaching movements are similar in both rodent and humans. Moreover, reaching movement abnormalities detected in rodent models of human neurodegenerative disease show similarities to human patients. Thus, skilled reaching tasks are a useful model for pre-clinical studies in rodents and clinical studies in humans.

Rung Walking to Assess Skilled Fore- and Hind Limb Function

Recent investigations showed that rats use skilled hind limb movements to adapt their gait pattern to a difficult territory. The rung walking task is a simple and sensitive test to simultaneously assess skilled fore- and hind limb movements [2]. The rung walking task resembles a horizontal ladder with rungs that can be adjusted individually. A regular pattern of rungs allows animals to anticipate rung location and learn a specific sequence of patterns across repeated test sessions. An irregular pattern prevents animals from learning the rung sequences and patterns can be changed from session to session.

Rung walking performance can be video recorded for further analysis of end point measures, such as number of placement errors, and qualitative analysis of the type of error made. As a result of limb-placement errors occurring in the affected limbs, the number of errors also increases in the intact limbs, thus revealing compensatory adjustments. Animals with lesions to various motor areas, such as motor cortex and its efferent corticospinal tract, and basal ganglia, show chronic impairments in limb placement. Furthermore, the rung walking task also monitors changes in errors in response to physiological variables such as aging and stress [3].

Conclusion

The design of standardized skilled movement tasks provides a reliable and high-resolution test strategy for studies of functional recovery. Repeated testing allows monitoring discrete changes in movement ability over time. The combination of quantitative and qualitative measures in skilled movement tasks elaborates recovery versus compensation, thus providing insights into underlying structural changes of regeneration and plasticity.

References

1. Metz, G.A., Whishaw, I.Q. (2000) Skilled reaching an action pattern: stability in rat (Rattus norvegicus) grasping movements as a function of changing food pellet size. *Behavioral Brain Research* **116**: 111-122.
2. Metz, G.A., Whishaw, I.Q. (2002) Cortical and subcortical lesions impair skilled walking in the ladder rung walking test: a new task to evaluate fore- and hindlimb stepping, placing, and co-ordination. *Journal of Neuroscience Methods* **115**: 169-179.
3. Metz, G.A., Schwab, M.E., Welzl, H. (2001) The effects of acute and chronic stress on motor and sensory performance in male Lewis rats. *Physiology and Behavior* **72**: 29-35.
4. Whishaw, I.Q. (2000) Loss of the innate cortical engram for action patterns used in skilled reaching and the development of

behavioral compensation following motor cortex lesions in the rat. *Neuropharmacology* **39**: 788-805.

5. Whishaw, I.Q. (2005) Prehension. In: *The Behavior of the Laboratory Rat* (Whishaw, I.Q., Kolb, B., eds.), 162-170.
6. Whishaw, I.Q., Suchowersky, O., Davis, L., Sarna, J., Metz, G.A., Pellis, S.M. (2002) Impairment of pronation, supination, and body co-ordination in reach-to-grasp tasks in human Parkinson's disease (PD) reveals homology to deficits in animal models. *Behavioral Brain Research* **133**: 165-176.

Measuring recovery of forelimb function after CNS injury in rodents and primates, with notes on man

Michael S. Beattie and Karen-Amanda Irvine
Brain and Spinal Injury Center, Dept. of Neurological Surgery, University of California, San Francisco,USA, michael.beattie@ucsf.edu

The development of reliable, valid, and quantifiable measurement of neurological function is important for the testing and translation of strategies aimed at improving function after CNS injury or disease. Tests of neurological function in spinal cord and brain injury range from descriptive locomotor measures to high speed kinematics. In this section of our symposium, we will describe a variety of tests aimed at determining loss and recovery of forelimb function in rats and monkeys, and will refer to techniques used to measure function and functional recovery after stroke and brain injury in man. The goal for the future is to determine common tests or common features of recovery that can then be used to integrate animal models with planning for outcome measures in human clinical trials of treatments for CNS injury.

A number of laboratories study recovery of function after unilateral cervical injuries in the rat, usually after partial or complete hemisections at high cervical levels. We and others have also devised unilateral and bilateral contusion lesion of the cervical cord in rats (Gensel et al, 2006). The unilateral lesions allow for contralateral comparisons in function, and also provide a useful model that has much less disability that either a bilateral cervical or thoracic contusion lesions. A variety of forelimb tasks have been adapted to this model (see Gensel et al, 2006), including the use cylinder test for limb use, a grooming test that engages a stereotyped forelimb motor pattern and can be scored, and locomotor analyses on the CatWalk apparatus (Hamers et al, 2006) such as the area occupied by the forepaw. These will be described and demonstrated in rats with unilateral cervical contusion injuries. A new test developed in the laboratory will also be described and demonstrated: rats are videotaped while eating various shaped pieces of cereal using both forepaws. They exhibit considerable dexterity in this task, and after injury, lose function of the ispilateral paw. Forepaw dexterity recovers in a predictable patter over time, and plateaus at different levels of function depending upon the severity of the contusion injury. This can be scored on a 10 point scale, Recovery curves are comparable to those seen in the Gensel et al (2006) study for the cylinder ets and the grooming test.

Additional tests of skilled limb movements will be described in the presentation by Dr. Metz.

In collaboration with the California Primate Spinal Cord Consortium, led by Drs. Mark Tuszynski and Reggie Edgerton, our laboratory has been involved in developing relatively simple and repeatable tests of forelimb function in rhesus macaques that have focal, unilateral, spinal cord lesions. These tests include a reach-grasp-and pull task, a food retrieval task, and a drawer-pull task that requires thumb-finger apposition for success. These tasks are also compatible with electromyographic (EMG) measurements that can demonstrate the degree of flexor-extensor co-activation. Parallels between these tasks and those used for the rats will be discussed.

Finally, a brief discussion of current literature describing the ways in which human stroke and traumatic brain injury patients are tested for arm and hand function and recovery will be given, with parallels and potential applications to the animal model techniques also discussed.

Supported by NIH grants NS-31193 and NS-38079, and the Roman Reed Fund (State of California).

References

1. Gensel, J.C., Tovar, C.A., Hamers, F.T.P., Deibert, R.C., Beattie, M.S., Bresnahan, J.C. (2006) Behavioral and histological characterization of unilateral cervical spinal cord contusion injury in rats. *Journal of Neurotrauma*, **23**, 36-54.
2. Hamers, F.P.T., Koopmans, C.G. and Joosten, E.A.J. (2006) CatWalk-assisted gait analysis in the assessment of spinal cord injury. *Journal of Neurotrauma* **23**, 537-548.

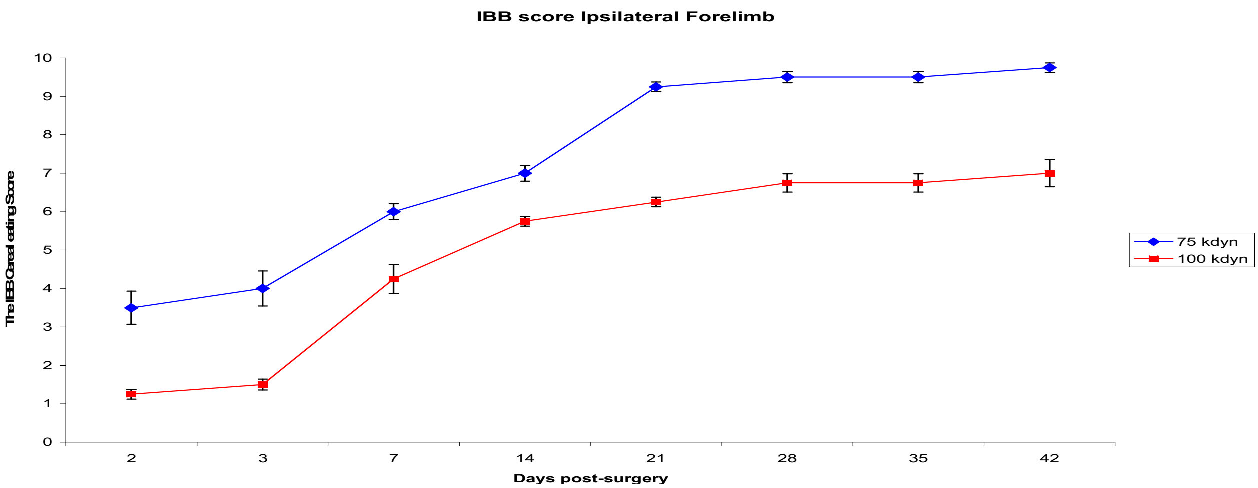


Figure 1. *Recovery of forelimb use during a cereal eating task as scored on the new rating scale after a 100 Kdyne (red) and 75 Kdyne (blue) contusion injury with the IH device. Weeks post-injury are shown across the bottom.*

Translational measures of behavioral function after spinal cord injury: A multivariate study of outcomes across species

A.R. Ferguson[1], G.C. Courtine[2,3], E.S. Rosenzweig[4], D.L. Jindrich[3,5], J.C. Gensel[6], K.-A. Irvine[1], V.R. Edgerton[3], M.H. Tuszynski[4], J.C. Bresnahan[1], and M.S. Beattie[1]

[1]*Dept of Neurological Surgery, University of California, San Francisco, San Francisco, CA, USA, adam.ferguson@ucsf.edu*
[2]*University of Zurich* [3]*Dept of Physiological Science, University of California, Los Angeles* [4]*Dept of Neurosciences, University of California, San Diego* [5]*Dept of Kinesiology, Arizona State University,* [6]*Dept Neuroscience, The Ohio State University, USA*

The past 20 years have seen significant progress in our understanding of the pathophysiology of spinal cord injury (SCI). As highlighted by the other presenters in this session, advances in our ability to measure behavioral changes in animal models have made a significant contribution toward this progress. In recent years a number of authors have developed new sensitive and reliable measures that provide insights into a range of functional states including locomotor function, forelimb function, autonomic function and sensory function after SCI. In addition, the rapid expansion of our understanding of the basic neuroscience of central nervous system trauma and plasticity has provided unprecedented opportunities to identify and target the mechanisms of injury with new therapeutics. Yet, to date, there have been few examples of new therapies that have made the translational leap from basic neuroscientific research to clinical application.

These translational difficulties may be due, in part, to the intrinsic complexity of SCI coupled with limitations of the analytical techniques that are commonly used to deal with SCI data. To understand the biological processes involved in SCI, researchers have developed animal models that mimic functional changes experienced by humans. However, the value of animal models depends, in part, on their ability to emulate the constellation of biological and functional changes observed in human SCI. Currently, there are many sensitive measures of behavioral function in animal models (Figure 1), however there is no unified strategy for integrating these functional measures to gain insights into the mechanisms underlying alteration and recovery of function following a SCI. One issue is the lack of universal standards for data acquisition, data formatting, and data sharing within the basic research community. This lack of standards reduces the replicability across laboratories and across species, and limits the translational potential of SCI findings. A second major problem is that quantitative integration of behavioral measures requires sophisticated statistical techniques that are not commonly used within the animal SCI literature.

In the prevailing approach, researchers attempt to understand the relationship between biology and functional behavior using simple correlation or stepwise regression: one functional measure is correlated with one biological measure. This process is then repeated many times in an attempt to assess how different functional measures relate to different biological outcomes. Researchers often apply a similar analytic strategy in an attempt to understand relationships between different functional measures: one measure is correlated with another, and this process is repeated for all comparisons of interest. However, there are major limitations with this type of analysis. First, this approach does not allow one to assess the degree of redundancy among different functional tests. Second, it is insensitive to complex, systematic changes that exist across several inter-related outcome measures. However, it is these multivariate changes that define the *syndrome* and are the target of therapeutic intervention (Figure 2).

Multivariate statistics such as principal components analysis provide a powerful alternative approach. This family of analytical techniques detects clusters of association among many different outcome measures while remaining sensitive to detailed information provided by each of the individual tests. In the context of SCI, such multivariate clusters can help identify the relevant groups of outcomes for a particular question of interest (e.g., neuroprotection vs. plasticity) or the

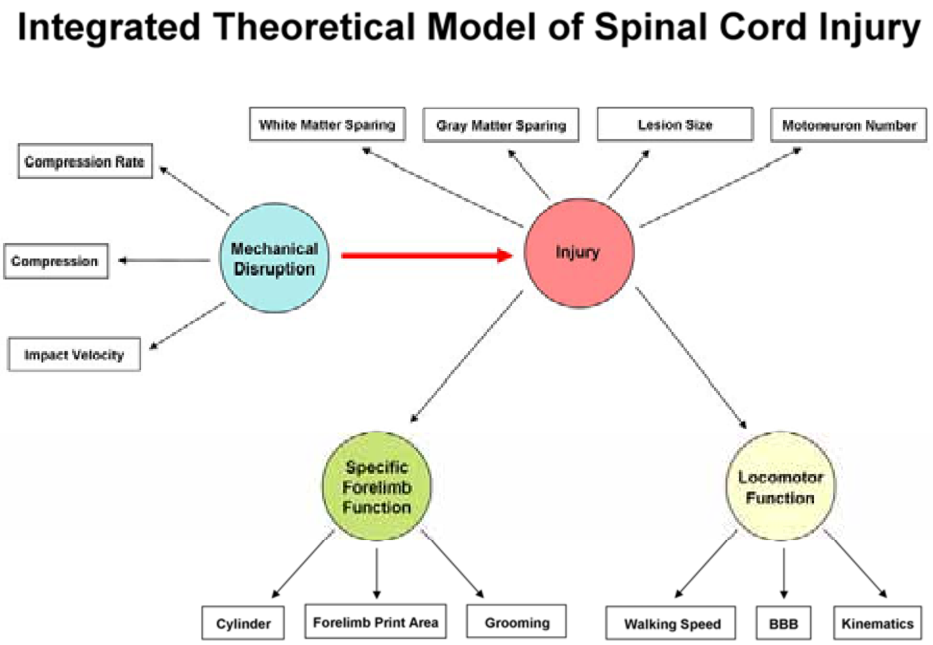


Figure 1. *Example of a simple integration scheme for a “typical” SCI experimental dataset. It is common for researchers to collect information within the same experimental subjects on the basic biomechanics of the injury, histological changes, and functional changes as measured on behavioral outcomes. However, these rich datasets are typically deconstructed with univariate statistical methods (e.g., correlation; ANOVA) which ignore information about the overlap among measures.*

outcomes that span several functional elements (Figure 2).

Here we illustrate the strength of this approach by applying multivariate techniques to identify, compare, and contrast syndrome-level features of cervical SCI in rodents and primates. The results identify a subset of outcomes that provide a unique window into the critical multivariate features of SCI in both rodents and primates. For example, within rat datasets, contusion injury affected a subset of behavioral measures (grooming, paw preference, print area an automated walkway) which had significant multivariate overlap with histology, forming a single multivariate factor (Factor 1). Other behavioral tests clustered together into different factors representing stability, locomotion, and horizontal ladder performance. These other factors were unrelated to tissue sparing, suggesting that these outcomes tap into subtle biological changes undetected by standard histology. As validation, we then tested the effect of graded cervical spinal cord injury using these multivariate features as the outcomes. The severity of SCI only had a significant effect on Factor 1, indicating that this cluster of measures is most sensitive to injury processes and should be used to evaluate therapeutic interventions that target these processes. Applying a similar approach to primate data, we were able to identify therapeutic effects that related to tissue damage, but were unrelated to functional recovery, suggesting that important therapeutic targets have yet to be discovered. Comparison of primate and rodent data revealed that a subset of tests showed temporal similarities across species in terms of behavioral losses and recovery. This suggests that SCI may engage analogous mechanisms in different species, and that it may be possible to select behavioral outcomes that are sensitive to translational changes. Further work is required to determine which measures represent translational outcomes and which reflect species-specific and model-specific changes. In the future, by targeting therapeutic interventions to translational multivariate features, SCI research may have improved success at discovering new therapeutic interventions that can be applied in a clinical setting.

Support: RO1 NS31193, NY State CoRE CO19772, F32 NS053059, RO1 NS49881, and the Roman Reed Fund

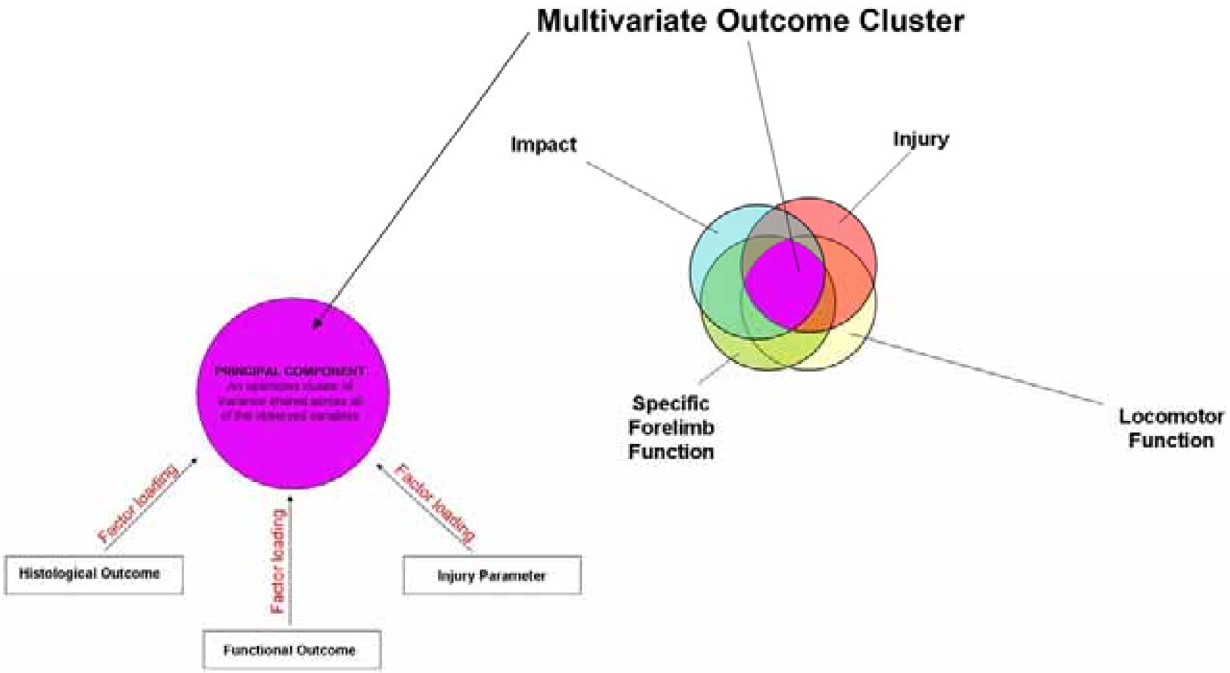


Figure 2. *Schematic illustration of the application of a multivariate method (in this case, principal components analysis) to identify clusters of outcomes within a typical SCI dataset. It is possible for clusters to reflect specific elements of SCI (e.g., impact biomechanics, tissue injury, or functional states) or overlap among all of the elements.*

Pain assessment Symposium

E.A. (Bert) Joosten
Head Exp.Anesthesiology , Academic Hospital Maastricht and Euron School for Neurosciences, University of Maastricht, the Netherlands, b.joosten@np.unimaas.nl

Currently, assessment of preclinical and clinical pain presents a unique problem compared to other health conditions, such as cancer or heart disease, which can be detected by objective biological measurements. Diagnosis of chronic pain depends upon subjective reports by patients on the presence and intensity of pain. However, comparable reports on sensory attributes cannot be obtained from laboratory animals without language skills. Nevertheless, attempts to assess chronic pain in non-human species have involved observations of spontaneous behavioral or physiological reactions to presumed sources of pain (Vierck CJ, Hansson PT and Yezierski RP (2008) Clinical and pre-clincial pain assessment: are we measuring the same thing? Pain 135, 7-10).

Basic research uses a variety of animal models, which often barely mimic the clinical syndromes. Clinical (pharmacological) trials are currently focused on etiological entities, whereas symptoms are hardly taken into account. In order to make progress in the research into pathophysiology of pain and development of new therapies, basic researchers as well as clinicians should take the whole sensory syndrome into consideration. During this symposium recent progress in the development and validation of objective technical methods to assess pre-clinical and clinical pain will be presented and discussed. The exiting development of new techniques and methods to assess pain can be expected to greatly improve discovery of therapeutic agents and the development of successful remedies for chronic pain.

Symposium contents

Measuring pain-induced gait adaptation: the CatWalk method
A.F. Gabriel, W.M.M. Honig, M.A.E. Marcus, E.A.J. Joosten

Traditional and novel behavioral models to evaluate pain and analgesia in preclinical studies
Thomas Christoph

Measuring pain
M. Cabanac

Facial expression to discriminate between pain and absence of pain in critically ill intubated adults during painful procedures
M. R. Arif and M.J. Grap

Measuring pain-induced gait adaptation: the CatWalk method

A.F. Gabriel, W.M.M. Honig, M.A.E. Marcus, and E.A.J. Joosten

Pain Management and research Center, Dept.of Anesthesiology, AZM, Maastricht, The Netherlands, a.gabriel@np.unimaas.nl

The Catwalk method is an automated quantitative gait analysis that allows the objective and rapid quantification of individual paw parameters as well as parameters related to interlimb coordination. Briefly, light from a fluorescent tube is sent through a glass plate. Light rays are completely reflected internally. As soon as anything, e.g. a rat's paw, is in contact with the glass surface, light is reflected downwards. It results in a sharp image of a bright paw print. Rats are trained in order to cross the pathway without hesitation at a constant speed. The complete run is recorded via a camera placed under the glass plate. Analysis of all gait parameters is performed through the use of an appropriate software.

The computer-assisted method of locomotor analysis was previously developed to assess motor changes on spinal cord injured rodents(Koopmans et al. 2005). However, based on the correlation of Catwalk parameters with the development of mechanical allodynia as assessed with the von Frey test (Vrinten and Hamers 2003), the present study aims at demonstrating that the CatWalk method allows an automated and complete measurement of gait related changes in a model of inflammatory pain; in other words, a pain-induced gait adaptation. In order to compare the behavioral outcomes from the CatWalk to a golden standard test, the von Frey test was used to assess the development of mechanical allodynia following the CAR-injection. Both test were performed first in the acute phase (0 – 48h post injection) and in the chronic phase (1 – 4 weeks) of the carrageenan model.

Carrageenan (CAR) is an irritant that induces an inflammatory response on injection into tissues. Injected into the rat knee, it results in oedema, pain and hyperalgesia of the injected limb. CAR-induced inflammation has been described as a model of acute inflammation through the first 24 h post injection and as a model of chronic inflammation after 1 week. In this study, both acute and chronic phases were studied.

Prior to injection ($t = -1$ day), CatWalk and von Frey tests were performed in order to determine the baseline. The experimental group was injected intra-articularly, at $t = 0$, with 2 mg of λ-carrageenan (200 µl) into the knee joint of the right hind paw. The control-group was injected intra-articularly with a saline solution (200 µl) into the same joint.

In the acute phase after injection of the CAR, von Frey test of both ipsilateral and contralateral hindpaws and CatWalk analysis were performed at $t = 2.5$, 4, 24 and 48 h post-injection. The CatWalk analysis of individual paw parameters like the intensity of the paw print or the time contact with the floor showed a significant effect after CAR injection into the knee. These CatWalk parameters were highly correlated with von Frey data. Furthermore, detailed CatWalk analysis of the gait (i.e. coordinated interaction between left and right hindlimb) showed a fully coordinated locomotion following the CAR injection, implicating that the accuracy of the rat gait is not affected. From these results we concluded that the CatWalk method allows an objective and detailed detection of a pain-induced gait adaptation, correlated to the development of mechanical allodynia in the acute phase of a CAR-induced knee inflammatory pain model.

Further validation of the CatWalk technique in measuring pain related changes was obtained with the use of a standard analgesic drug, Fentanyl. Twenty four hours after the CAR-injection, rats were subcutaneously injected with Fentanyl (25µg/kg). Control rats were injected with saline solution. CatWalk and von Frey tests were performed prior to experiment, prior to Fentanyl injection and at 20, 45 and 120 minutes post-Fentanyl injection.

Von Frey results showed a significant recovery of 88% from the carrageenan-induced mechanical allodynia at 20 minutes post-Fentanyl injection. CatWalk analysis (individual paw- and coordination-related parameters) showed a significant recovery (from 50 to 82%) of the carrageenan-induced gait changes at 20 and 45 minutes post-Fentanyl injection.

If compared to the von Frey results, CatWalk analysis allows a measurement of the analgesic effect of Fentanyl in an acute inflammatory pain model which is more sensitive (less variation and longer detection of the analgesic effect), more objective (fully computerized and therefore independent of the experimenter) and more reproducible (less variability between results and thus less animals needed). Herewith, we validated the use of the CatWalk in detection of pain related gait changes. Our results strongly favor the use of the CatWalk as a new tool for preclinical assessment of analgesic drugs.

Finally, the use of the Catwalk technique in the chronic phase of the CAR-induced inflammation model was studied. Von Frey test of both ipsilateral and contralateral hindpaws and CatWalk analysis were performed at $t = 1$, 7, 14, 21 and 28 days post-injection. Von Frey results showed a significant development of mechanical allodynia 4 hours after the injection of CAR in the rat knee, and a plateau like-phase was observed up to 21 days. CatWalk parameters related to individual paw (e.g. intensity of the paw printing) showed significant changes after 24 hours, identical to those reported in the acute phase. From 7 days post-injection on, no changes could be noted in any of the CatWalk parameters. For instance, a significant reduction in the duty factor of the paw print could be measured at day post-operation 1 (DPO1) in rats from the experimental group but changes do no longer exist at DPO7. Locomotion can be considered as a daily life activity, i.e. it is a voluntarily initiated movement, repeated several times a day. It is therefore likely to explain our results in the context of habituation. Daily use of the painful limb can lead to habituation which further results in a subsequent attenuation of the behavioral gait adaptation.

In conclusion, the use of the CatWalk as a method to assess a pain-induced gait adaptation in correlation with mechanical allodynia needs to be carefully restricted to identified experimental pain models. However, when appropriated, the CatWalk technique is a powerful tool that allows an objective and sensitive measurement of the pain-induced gait adaptation.

References

1. Koopmans GC, Deumens R, Honig WM, Hamers FP, Steinbusch HW, Joosten EA. 2005. The assessment of locomotor function in spinal cord injured rats: the importance of objective analysis of coordination. *J Neurotrauma* **22(2)**:214-225.
2. Vrinten DH, Hamers FF. 2003. 'CatWalk' automated quantitative gait analysis as a novel method to assess mechanical allodynia in the rat; a comparison with von Frey testing. *Pain* **102**(1-2):203-209.

Traditional and novel behavioral models to evaluate pain and analgesia in preclinical studies

Thomas Christoph

Preclinical Research, Grünenthal GmbH, Aachen, Germany, thomas.christoph@grunenthal.de

Pain is an unpleasant sensory and emotional experience associated with actual or potential tissue damage, or described in terms of such damage. Traditionally, animal pain research has focused on the sensory component of pain and a battery of models was developed. Sensory testing was performed on somatic structures which were readily accessible (e.g. hind paws and tails of rodents). However, over time research also focused on non somatic organs such as visceral hollow organs or nerves and led to models of visceral and neuropathic pain. Using acute or chronic inflammation introduced models of inflammatory pain. Recently, models which assess the affective component of pain have been developed.

While working with rodents allows for precise analgesic profiling of different molecular mechanisms in acute and chronic pain conditions, many confounding factors have been described which complicate the analysis of novel drug candidates. In addition, gender, species and even strain differences in nociception and analgesia have been described. Thus, the analysis of knock-out mice is often biased by the genetic background of the originally used inbred strains.

A variety of stimuli is used in sensory testing, and many variables contribute to the experimental outcome, such as stimulus quality, intensity and kinetics. While most studies determine thresholds, qualitative assessment of animal pain behavior is required to address symptoms such as hyperalgesia. Careful selection of controls is essential since many aspects of analgesia can only be uncovered when comparing the treatment effects to the appropriate control groups.

A special feature of pain research is the possibility to translate animal models into early phases of clinical research. Quantitative sensory testing techniques are using similar stimuli and read-outs in volunteers and/or pain patients. Thus, pain research offers a chance to optimize clinical drug candidates for proof of concept studies.

Measuring pain

M. Cabanac
Faculty of Medicine, Laval University, Quebec, Qc, Canada, michel.cabanac@phs.ulaval.ca

It is impossible to access other people mental space other than from the behavior they display or the verbal responses they tell. Even those signs may be misleading if the participants decide to deceive the experimenter and fake or lie. The student of pain in addition faces two other methodological hindrances:

a) because pain stimuli are harmful, e.g. pain is often produced in experiments with tying a tourniquet around an arm or leg to cut off circulation and starve downstream tissues of oxygen. The pain is acute but there is a risk to harm those participants who might not be able to feel pain, e.g. in congenital indifference to pain (see experiment 1);

b) the participants might be willing to feel pain for various other interfering motivations (see experiment 2) or mental diseases. The anoxia method for seeing how long you can withstand the pain is especially dangerous, for lesions may develop in the oxygen-starved tissues of highly motivated participants.

The general principle of conflicting motivations avoids largely most of these difficulties and may be used with animals as well (see experiment 3).

Experiment 1, was simple psychophysics using not anoxia or electrical stimuli but harmless simple temperature stimuli. The participants were asked to report and quantify the intensity of pleasure/displeasure experienced when dipping one hand in well stirred water at a regulated temperature betwwen 10°C and 50°C. As the pain thresholds are 15°C and 45°C the range for pain is largely available without being dangerous or harmful. That method was applied to normal healthy participants and to two persons with congenital infifference to pain. The results showed that within that limits the sensation covered the whole range between very unpleasant and very pleasant, according to the participants internal core temperature.

Experiment 2. Isometric muscle contraction, another method to inflict pain, was used in Experiment 2. The participants had to adopt a sitting position with their backs against a wall and their lower limbs at right angles, but without a seat. Such a position can be maintained only by keeping extensor muscles fully contracted[1]. The participant doesn't move and the tonic contraction keeps blood from irrigating the thigh muscles. A painful ischemia soon appears and steadily increases to the point of becoming intolerable. This was the pain that participants endured in conflict for a monetary reward, as in that experiment they received money at various rates over different sessions, The longer they tolerated pain, the higher the amount of money they recieved at the end of the session. The results showed that it was thus possible to calibrate pain against money as the duration tolerated was proportional to the logarithm of money earned.

Experiment 3. In that experiment both the methods used in Experiment 1: temperature stimulus, and 2: conflict of motivations, were used with rats. The animals wer housed in a warm 'home' with food and water available *ad libitum*, and they had access to a palatable reward at the end of a 16 m long zig-zag alley. But the ambient temperature outside the warm home was -15°C, *i.e.* cold enough to presumably produce pain in these rats. The results showed that the rats repeatedly ran to the palatable taste stimuli but at the cost of pain as they had tail tip and ear tip frostbite necrosis. Yet the pain must not have been severe as the animals were not forced to endure it, they could have stayed in their warm home, but decided by themselves that it was worth to obtain the palatable reward[2].

Thus, these methods avoid most of the difficulties and dangers encountered in quantitative experimental studies of pain.

[1]This exercise is part of the training for competing downhill skiers.

[2]That experiment had been approved by Laval University "Comité de protection des animaux."

Facial expression to discriminate between pain and absence of pain in critically ill intubated adults during painful procedures

M. R. Arif and M.J. Grap
School of Nursing, Virginia Commonwealth University, Richmond, Virginia, marif@mcvh-vcu.edu

Purpose

To describe the evidence of behavioral and biological markers related to facial expression in non-communicative critically ill patients experiencing pain during procedures.

Research Questions

What is the empiric evidence of the relationship of facial expressions to the experience of pain in the critically ill?

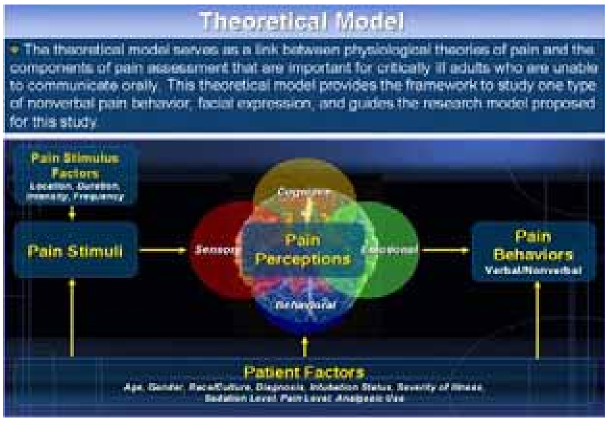


Figure 1. *Theoretical model*

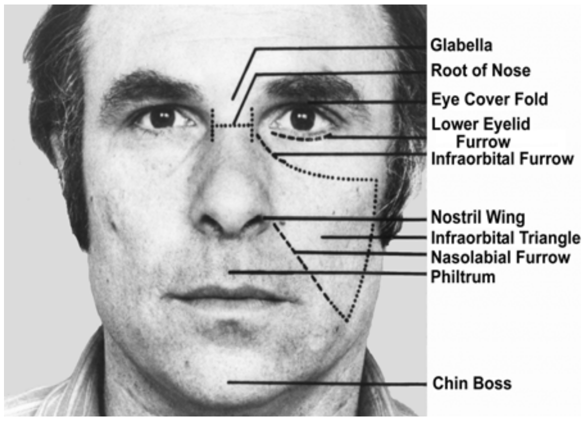
Background

Pain is a complex multidimensional concept that is difficult to define. Individual pain experiences influence cognitive, emotional, and behavioral responses (Figure 1). Pain assessment is a significant challenge in critically ill adults, especially those who are unable to communicate their pain level. In critical care many factors may alter oral communication including tracheal intubation, reduced level of consciousness, use of restraints, sedation, and paralyzing drugs. These patients are more likely to receive inadequate analgesics than those who are able to communicate, but just as likely to experience painful illnesses. The need for optimal pain assessment in adult critical care settings is essential since it has been reported that 35% to 55% nurses underrate the patient's pain [1-3]. and in one study 63.6% of patients did not receive any medications before and/or during painful procedures [4].

Furthermore, unconscious or sedated patients cannot communicate their level of pain using numeric pain rating scales (NRS) (0-to-10) and are therefore at risk for being inadequately medicated for pain [5-6]. Inaccurate pain assessments and resulting inadequate treatment of pain in critically ill adults can lead to significant physiologic consequences such as increases myocardial workload which can lead to myocardial ischemia, or impairs gas exchange causing a cascade of events which can lead to pneumonia [7]. The first step in providing adequate pain relief for patients is systematic and consistent assessment and documentation of pain. Identification of the optimal pain scales for non-communicative patients have been the focus of several studies. To date, however, no one pain assessment tool is universally accepted for use in the non-communicative patient. A common component of behavioral pain tools is evaluation of

Figure 2. *Using The Observer XT 7.0 to code facial action units*

facial behaviors. Although use of facial expression is an important behavioral measure of pain intensity, precise and accurate methods for interpreting the facial expression of pain has not been empirically evaluated in critically ill, non-communicative patients.

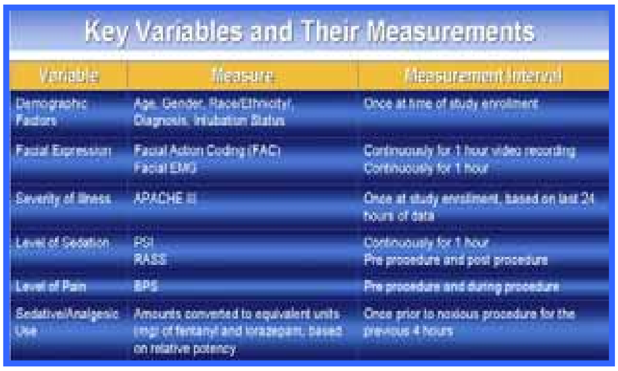


Findings

The face reveals a wealth of information about human behavior. The most frequently used pain behavior in pain evaluation scales for patients who cannot orally communicate is facial expression. Facial expression has been studied for centuries, dating back to Charles Darwin's "The Expression of Emotions in Man and Animals"[8] reporting observations of why particular facial expressions occur with particular emotions. Ekman and Friesen, pioneers of the Facial Action Coding System (FACS), identified 6 universal facial expressions [9]. They demonstrated that observers judgments of anger, disgust, fear, sadness, happiness and surprise made by preliterate people as isolated as New Guineans were no different than judgments made by college students in eight literate cultures around the world [10]. Patients who are unable to orally communicate due to altered mental status, sedation, unconsciousness, or cognitive impairment have been observed to show facial expression during painful stimuli [11-12]. Facial expressions provide an important behavioral measure for the study of emotion, cognitive processes, and social interaction. Use of facial expression is also an important behavioral measure of pain intensity, but precise and accurate methods for interpreting facial expressions of pain in non-communicative critically ill adults has not been identified.

Facial expression in pain

Facial expression specific to pain has been studied using the Facial Action Coding System (FACS). These include lowered brows, raised cheeks, tightened eyelids, a raised upper lip or opened mouth, and closed eyes. Craig and Patrick [13] used the FACS to identify six facial activity associated with exposure to noxious stimulus (AUs 6-7-cheek raise, lids tight; 10-upper lip raise;12-lip corner pull; 25-lips part; 26-27-jaw drop-mouth stretch; and 43-45-eyes closed-blink). Prkachin [14] focused on pain behavior of adults during three pain stimulus (electric shock, cold, pressure, and muscle ischemia). He found 4 actions evident with pain, increasing in intensity/duration: brow lowering (AU4), tightening and closing of the eye lids (AU6/AU7), and nose wrinkling/ upper lipraising (AU9/AU10).



Key Variables and Their Measurements

Variable	Measure	Measurement Interval
Demographic Factors	Age, Gender, Race/Ethnicity/, Diagnosis, Intubation Status	Once at time of study enrollment
Facial Expression	Facial Action Coding (FAC) Facial EMG	Continuously for 1 hour video recording Continuously for 1 hour
Severity of Illness	APACHE III	Once at study enrollment, based on last 24 hours of data
Level of Sedation	PSI RASS	Continuously for 1 hour Pre procedure and post procedure
Level of Pain	BPS	Pre procedure and during procedure
Sedative/Analgesic Use	Amounts converted to equivalent units (mg) of fentanyl and lorazepam, based on relative potency	Once prior to noxious procedure for the previous 4 hours

Figure 3. *Key variables and measurements*

Methods

This descriptive research design will be conducted in 933-bed tertiary care university medical center. Sample of 100 subjects will be recruited Surgical Trauma ICU and Medical Respiratory ICU who are non-communicative. Key variables and their measurements are listed in Figure 3

Facial Expression will be measured using the Facial Action Coding System with slow action video and stop-frame feedback. The basic elements of FACS are 44 action units (AUs). Each AU represents the movement of a single facial muscle or a group of muscles, which move as a unit. FACS has been integrated in a recent new application of scoring facial expressions with the use of The Observer XT 7.0 software (see Figure 2). This The Observer XT allows for synchronized recording and playback of four full-resolution media files, physiological data, and EMG.

Facial muscle activity will be measured using electromyography (EMG). The EMG signals will be recorded and processed by a MyoSystem 1200 (Noraxon, Inc., Scottsdale, AZ) instrument. EMG recordings are beneficial because of their higher temporal resolution and the ability to register EMG activity automatically and continuous over an extended period of time, as well as the objective evaluation of EMG signals.

Discussion

Facial expressions provide an important behavioral measure for the study of emotion, cognitive processes, and social interaction. Use of facial expression is an important behavioral measure of pain intensity, but precise and accurate methods for interpreting facial expressions of pain in non-communicative critically ill adults has not been identified. The goal of this presentation is demonstrate use of biobehavioral technology such The Observer XT tool for recording and analyzing behaviors measured by FACS, EMG, and other physiological data.

References

1. Hamill-Ruth, R. J. & Marohn, M. L. (1999). Evaluation of pain in the critically ill patient. *Crit Care Clin.,* **15**, 35-vi.
2. Puntillo, K. (1997). Stitch, stitch..creating an effective pain management program for critically ill patients. *AmJ.Crit Care, 6,* 259-260.
3. Puntillo, K. A. (1990). Pain experiences of intensive care unit patients. *Heart Lung,* **19**, 526-533.
4. Puntillo, K. A., Wild, L. R., Morris, A. B., Stanik-Hutt, J., Thompson, C. L., & White, C. (2002). Practices and predictors of analgesic interventions for adults undergoing painful procedures. *Am.J.Crit Care,* **11**, 415-429.
5. Hall-Lord, M. L., Larsson, G., & Steen, B. (1998). Pain and distress among elderly intensive care unit patients: comparison of patients' experiences and nurses' assessments. *Heart Lung, 27,* 123-132.
6. Ferguson, J., Gilroy, D., & Puntillo, K. (1997). Dimensions of pain and analgesic administration associated with coronary artery bypass grafting in an Australian intensive care unit. *J.Adv.Nurs.,* **26**, 1065-1072.
7. McArdle P. (1999). Intravenous analgesia. Crit Care Clin. 14, 89-104.
8. Darwin, C. (1998). The expression of the emotions in man and animals / Charles Darwin ; with an introduction, afterword, and commentaries by Paul Ekman. (3 ed.) New York: Oxford University Press.
9. Ekman, P. & Friesen, W. V. (1978). *Investigator's guide to the Facial Action Coding System.* Palo Alto, CA: Consulting Psychologists Press.
10. Ekman, P., Friesen, W. V., O'Sullivan, M., Chan, A., acoyanni-Tarlatzis, I., Heider, K. et al. (1987). Universals and cultural differences in the judgments of facial expressions of emotion. *J.Pers.Soc.Psychol.,* **53**, 712-717.
11. Puntillo, K. A., Miaskowski, C., Kehrle, K., Stannard, D., Gleeson, S., & Nye, P. (1997). Relationship between behavioral and physiological indicators of pain, critical care patients' self-reports of pain, and opioid administration. *Crit Care Med.,* **25**, 1159-1166.
12. Puntillo, K. A., Morris, A. B., Thompson, C. L., Stanik-Hutt, J., White, C. A., & Wild, L. R. (2004). Pain behaviors observed during six common procedures: results from Thunder Project II. *Crit Care Med.,* **32**, 421-427.
13. Craig, K. D. & Patrick, C. J. (1985). Facial expression during induced pain. *J.Pers.Soc.Psychol.,* **48**, 1080-1091.
14. Prkachin, K. M. (1992). The consistency of facial expressions of pain: a comparison across modalities. *Pain,* **51**, 297-306.

Behavioral testing of zebrafish Symposium

Robert Gerlai
University of Toronto, Mississauga, Ontario, Canada, rgerlai@utm.utoronto.ca

Zebrafish has been in the forefront of developmental biology for the past three decades and as a result numerous genetic tools have been developed for this species. Briefly, zebrafish has become the species of choice for geneticists along with the classical species of genetics, the fruit fly and the house mouse. Zebrafish, however, enjoys some advantages over these latter species that may make it appropriate particularly for translational research. Although it is small and almost as easy and cheap to keep in the lab as the fruit fly, it is a vertebrate species with anatomy, physiology, and brain function characteristics very similar to those of other vertebrates including our own. The high nucleotide sequence homology between zebrafish and human genes coupled with the ease of maintenance and prolific nature of this species will make zebrafish one of the most preferred model organisms in biomedical research. However, as far as understanding of its brain function and behavior is concerned, zebrafish is a rather novel organism. The bottleneck in behavior genetics and brain research with zebrafish is exactly this: without sophisticated behavioral methods and good understanding of brain function, analysis of the biology and genetics of brain and behavior function is difficult. The current symposium brings together four speakers from diverse areas of behavioral neuroscience and shows how zebrafish may be utilized. The pioneering work presented in the symposium suggests that zebrafish will indeed be useful in the analyses of diverse brain functions. Dr. Karlsson will talk about how one can measure sleep in different organisms and how zebrafish may advance our understanding of this important but still understudied brain function. Dr. Willemsen will discuss how zebrafish may be employed to investigate and model Parkinson’s disease. Dr. Norton will present approaches as to how one can analyze mood disorders with zebrafish and Dr. Gerlai will show how automated behavioral test paradigms may be developed for the investigation of the effects of acute and chronic alcohol administration in zebrafish.

Symposium contents

Measuring sleep in complex and simple organisms
K.Æ. Karlsson

Zebrafish as a new model organism for Parkinson’s disease
Rob Willemsen, Wiebren Hasselaar, Herma van der Linde, Vincenzo Bonifati

Zebrafish: Development of automated behavioral test paradigms
Robert Gerlai

Approaches to analyse mood disorders in zebrafish
W.H.J. Norton, K. Webb, M. Harris, N. Rohner, C. Nüsslein-Volhard, J. Ninkovic, A. Folchert, and L. Bally-Cuif

Measuring sleep in complex and simple organisms

K.Æ. Karlsson
Department of Biomedical Engineering, School of Science and Engineering, Reykjavik University, Reykjavik, Iceland, karlsson@ru.is

Recently, surprising statistical regularities have been revealed in the structure of bouts of sleep and wakefulness [1]. This characterization of sleep offers a novel method of measuring and classifying behavioral states. In contrast to the standard in the field, this method renders sleep comparable across phylogeny and ontogeny and, thus, opens new ways of dissecting sleep at the behavioral, pharmacological, and genetic levels.

In spite of being poorly understood phenomenon, sleep comes equipped with virtually undisputed definition, in which sleep is measured and defined as a constellation of physiological events – centered on the electroencephalogram. According to this tradition sleep is measured in 30 second bins and depending on the amplitude and frequency of the electroencephalogram, the amplitude of the electromyogram, and the presence or absence of eye movements, each 30 second epoch is assigned a state: Wake, rapid-eye-movement sleep, or slow-wave-sleep 1-4 [2]. However, strict adherence to this approach, while clinically useful, virtually defines sleep out of existence large in groups of animals and age groups. Invertebrates, such as the fruit fly, which clearly exhibit all behavioral symptoms of sleep cannot be considered sleeping under this traditional method [3]. Moreover, neonates cannot either. A human newborn quietly sleeping at its mother's bosom, the duckling under its mother's wing, or a rat pup huddled against its littermates; these are prototypical examples of sleep. Yet, neither the human neonate, duckling, nor rat pup exhibit all the indices that have become the gold standard for defining sleep. Sleep research, thus, has been conducted almost exclusively in adults; and only in a handful of mammalian species. Sleep was first measured and described in adults [4], the dominant terminology of the field is derived from work done with adults [2], and the neural substrates of sleep have been elucidated from work done with adults [5]. (For an exeption, see [6]). Accordingly, most theories of sleep can only be applied to infants of our own species, or to other species, with great difficulty [7].

In adult humans the duration of sleep bouts exhibit an exponential distribution with the rule $P(t) \sim \exp(-t/\tau)$ where t is an individual sleep bout, whereas, wake bouts exhibit a power-law distribution with the rule $P(t) \sim t{-}\alpha$ where t is an individual wake bout [1]. Subsequently, it was demonstrated that the wake bouts exhibit a scale-free power law behavior with an exponent, α, that remains constant across species (humans, cats, rats, and mice). In contrast, sleep bout durations follow an exponential distribution where τ represents a characteristic time scale whose main determinants are body size and metabolic rate [8]. In neonatal rats, both sleep and wake bouts exhibit exponential distribution immediately after birth, with a clear power-law behavior of wake bouts emerging only after the second postnatal week; this occurs in spite of very little change in the overall duration of wake bouts; τ, on the other hand, increases with age [9]. Thus, the power-law exponent α is constant across multiple adult species, but switches from exponential to power-law behavior during development. In contrast, the sleep-related time constant τ varies across species and age. Importantly, the only information needed to calculate α and τ is sleep and wake durations; one does not need detailed information about the transitions *between* sleep states or information about events within a given state. Given the right measuring system, this method can be employed to the simple, genetically tractable zebrafish [10]. These efforts could render sleep in humans and zebrafish meaningfully comparable and open new venues in sleep research.

The values, α and τ, thus, represent novel way of thinking about sleep, and offer a novel, simpler method of characterizing sleep states; a method that is based on the stability of behavioral states. This novel method of characterizing sleep states could become the new standard of classifying sleep, its disorders, and its development through the life-span.

References

1. Lo, C.C., Amaral, L.A.N., Havlin, S., Ivanov, P., Penzel, T., Peter, J.H., Stanley, H.E., (2002) Dynamics of sleep-wake transitions during sleep. *Europhysics Letters* **57**(5): p. 625-631.
2. Rechtscaffen, A. and A. Kales, eds. (1968) A Manual of Standardized Terminology, Techniques and Scoring System For Sleep Stages of Human Subjects. Public Health Service, US Government Printing Office: Washington, D.C.
3. Shaw, P.J., et al. (2000) Correlates of sleep and waking in *Drosophila melanogaster*. *Science* **287**(5459): p. 1834-7.
4. Dement, W. and N. Kleitman. (1957) The relation of eye movements during sleep to dream activity: an objective method for the study of dreaming. *Journal of experimental psychology* **53**: p. 339-346.
5. Pace-Schott, E.F. and J.A. Hobson. (2002) The neurobiology of sleep: genetics, cellular physiology and subcortical networks. *Nature Reviews Neuroscience* **3**(8): p. 591-605.
6. Karlsson, K.Æ., et al. (2005) The neural substrates of infant sleep in rats. *PLoS Biology* **3**(5): p. 891-901.
7. Rechtschaffen, A. (1998) Current perspectives on the function of sleep. *Perspectives in Biology and Medicine* **41**(3): p. 359-390.
8. Lo, C.C., et al. (2004) Common scale-invariant patterns of sleep-wake transitions across mammalian species. *Proc Natl Acad Sci U S A* **101**(50): p. 17545-8.
9. Blumberg, M.S., et al. (2005) Dynamics of sleep-wake cyclicity in developing rats. *Proc Natl Acad Sci U S A* **102**(41): p. 14860-4.
10. Zhdanova, I.V. (2006) Sleep in Zebrafish. *Zebrafish* **3**(2): p. 225.

Zebrafish as a new model organism for Parkinson's disease

Rob Willemsen, Wiebren Hasselaar, Herma van der Linde, Vincenzo Bonifati
Dept. of Clinical Genetics, Erasmus MC, Rotterdam, The Netherlands, r.willemsen@erasmusmc.nl

Parkinson's Disease

Parkinson's Disease (PD) is the second most common human neurodegenerative disorder. PD patients present aberrant motor activity with resting tremor, muscular rigidity, bradykinesia and postural imbalance [1, 2]. The cause of these symptoms is loss of dopaminergic neurons in the substantia nigra, pars compacta. Another pathological hallmark of PD is the frequent formation of Lewy bodies in surviving neurons. Lewy bodies are cytoplasmic aggregates of insoluble proteins. About 85-90% of the PD cases are sporadic with a complex aetiology. In 10-15% of the PD patients, a familial predisposition can be found. Linkage analysis in these families conclusively identified α-synuclein, parkin, *PINK1, DJ-1*, *LRRK2* and *ATP13A2* as PD-causing genes [1, 3, 4]. The normal cellular function of these genes suggests that several pathways can be involved in neurodegeneration, including oxidative stress, proteosomal and mitochondrial dysfunction and protein aggregation and misfolding [5]. Our research is focused on two PD-causing genes, that is, *LRRK2* and *ATP13A2*. *LRRK2* has been implicated in both familial and sporadic cases but the precise role of the LRRK2 protein in neurodegeneration is not clear. Clinical relevance is implicated by the similarity between the PD-phenotype of patients carrying mutations in this gene and that of sporadic PD patients. The relatively high mutation rate in sporadic PD in some populations also adds to the hypothesis that LRRK2 plays a key role in PD aetiology [6]. The other gene, ATP13A2, was identified in Kufor-Rakeb Syndrome. Patients are bound to bed before the age of 15 due to severe Levodopa-responsive Parkinsonism and pyramidal symptoms. This gene was selected because of its severe phenotype and autosomal recessive inheritance. In the human patients these mutations result in a loss of functional protein. The co-localization of ATP13A2 and β-Synuclein [7] in lysosomal inclusions also suggest a role for this gene in Lewy Body formation.

Zebrafish

Recently we have implemented a new simple vertebrate model with a richness of embryology/anatomy and genetics knowledges, creating a new vertebrate model system for PD, the zebrafish (*Danio rerio*).

The zebrafish, a fresh water tropical fish, is a premiere model organism to study vertebrate development. Fast external development and transparency during embryogenesis allow for visual screening at the macroscopical and microscopical level, including visualization of organogenesis. High fecundity and short generation times facilitate genetic analyses. Importantly, The Sanger Institute has recently released assembly version Zv7, an almost finished sequence, of the 1.5Gb zebrafish genome. Zebrafish may be a particular powerful model for the study of human disease because many cellular processes are conserved throughout vertebrate evolution, including the corresponding disease genes. Finally, the ability to generate

transgenic zebrafish allows for (over-) expression as well as for the suppression of gene expression during early development (using morpholino gene knockdown techniques) creates easy access to new animal models. Although the zebrafish brain does not contain a mesencephalic region comparable to the substantia nigra, treatment with MPTP (a PD-inducing drug in humans, apes and to a lesser extent mice) showed a direct effect on diencephalic dopaminergic neurons. This resulted in a loss of diencephalic dopaminergic neurons and an aberrant swimming pattern, illustrating that zebrafish can develop a phenotype comparable to PD [8]. Importantly, the homologues of the two selected PD-causing genes have been identified in zebrafish and morpholinos to study gene knockdown are available.

Aims

Our research project aims at accomplishing two objectives. The first objective is a fundamental research question: Understanding the cellular function of atp13a2 and lrrk2 during brain development and in adult brain.

The second objective is of translational nature: understanding the molecular mechanisms underlying PD by generating animal models of PD using zebrafish.

Methods

Manipulation of gene expression in zebrafish includes: gene knockdown and transgenic strategies. For gene knockdown, we will use antisense oligonucleotides with a synthetic backbone called morpholinos (MOs). MOs are 25 basepairs long and are designed to uniquely bind to our gene of interest, thereby disrupting translation initiation or pre-mRNA splicing by binding to the pre-mRNA. MOs are used for functional genomic applications and have been shown to be very effective in zebrafish. MOs are injected in the yolk sac of 1 or 2 cells-stage embryos and knockdown of gene function is transient (till 5 days post fertilization=dpf). For the generation of transgenic zebrafish, we will microinject plasmid DNA,

containing our gene of interest fused to EGFP under the control of a brain-specific promoter (GATA-2 or Tyrosine Hydroxylase (TH) promoter) into the cytoplasm of a 1-cell stage embryo.

Both knockdown morphants and transgenic zebrafish will be further characterized for the presence of a phenotype, including microscopical screening for morphological features, in situ hybridization (ISH) strategies using molecular probes to study gene expression patterns, immunohistochemical techniques (i.e. TH expression) and biochemical analyses (Western blot). Another important aspect of characterization of a phenotype is the use of behavioural tests. Recently, a system to monitor locomotor behaviour in zebrafish has been established (Noldus Information Technology bv; EthoVision). This system enables to record movement and swimming patterns of zebrafish embryos (6 dpf) in a highthroughput fashion. Currently, this system is implemented at Erasmus MC. Activity is monitored by a camera and the software detects the fish. This data visualizes the swimming pattern and allows for statistical analysis of locomotor activity (figure 1).

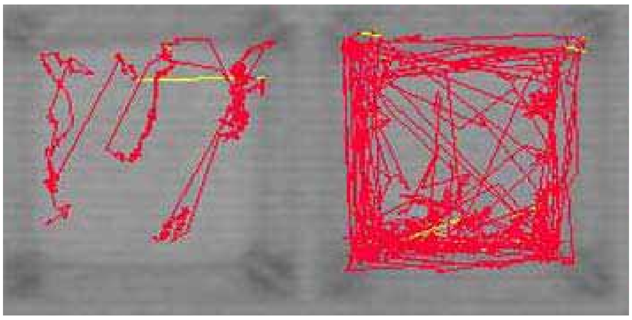


Figure 1. *Zoom in on 2 wells of a 96 well plate with plots of the swimming pattern of larvae at 6dpf during 5 minutes. The larva in the left well is treated with 10μM Diazepam, while the larva on the right is an untreated control*

References

1. Hardy, J., et al. (2006) Genetics of Parkinson's disease and parkinsonism. *Ann Neurol* **60(4)**, 389-98.
2. Twelves, D., K.S. Perkins, and C. Counsell. (2003) Systematic review of incidence studies of Parkinson's disease. *Mov Disord* **18(1),** 19-31.
3. Gasser, T. (2005) Genetics of Parkinson's disease. *Curr Opin Neurol*. **18(4)**, 363-9.
4. Ramirez, A., et al. (2006) Hereditary parkinsonism with dementia is caused by mutations in ATP13A2, encoding a lysosomal type 5 P-type ATPase. *Nat Genet* **38(10)**, 1184-91.
5. Wood-Kaczmar, A., S. Gandhi, and N.W. Wood (2006) Understanding the molecular causes of Parkinson's disease. *Trends Mol Med.* **12(11)**, 521-8.
6. Bonifati, V. (2007) LRRK2 low-penetrance mutations (Gly2019Ser) and risk alleles (Gly2385Arg)-linking familial and sporadic Parkinson's disease. *Neurochem Res* **32(10)**, 1700-8.
7. Wei, J., et al. (2007) Enhanced lysosomal pathology caused by beta-synuclein mutants linked to dementia with Lewy bodies. *J Biol Chem.* **282(39)**, 28904-14.
8. Panula, P., et al. (2006) Modulatory Neurotransmitter Systems and Behavior: Towards Zebrafish Models of Neurodegenerative Diseases. *Zebrafish* **3(2),** 235-47.

Zebrafish: Development of automated behavioral test paradigms

Robert Gerlai
Department of Psychology, University of Toronto @ Mississauga, Mississauga, Ontario, Canada, robert_gerlai@yahoo.com

Due to the arsenal of genetic techniques developed and the amount of genetic information accumulated for zebrafish over the past three decades, this species has been becoming popular in several disciplines of biology including behavioral neuroscience. The prolific nature, small size, and ease of maintenance of zebrafish make this species a potentially excellent organism for high throughput mutation or drug screening. However, compared to classical rodent species, we know much less about the behavioral characteristics of zebrafish and thus we often do not have appropriate behavioral test paradigms. In the present talk, I will show a range of simple behavioral applications with which we study social behavior, learning and memory, and alcohol induced behavioral changes in zebrafish.

We have developed a custom software application with which we quantify the distances among members of the zebrafish shoal. We discovered that although distances among adult zebrafish remain fairly constant over extended periods of observation time, shoal cohesion fluctuates with a characteristic amplitude and fast (5-10 sec) frequency. We also discovered that shoaling tendencies change during development with a characteristic increase of shoal cohesion between 20-40 days after hatching. We have also conducted numerous learning paradigms to analyze mnemonic and cognitive characteristics of zebrafish. We confirm that similarly to other cyprinids, zebrafish are capable of acquiring simple associative learning tasks (association between two stimuli) and more complex spatial learning tasks (association between reinforcement and a set of diffuse external cues). In addition, we have tested different reinforcers and found that the sight of conspecifics may be an excellent reward for zebrafish and may motivate fish in classical as well as instrumental conditioning paradigms. Based on this, we developed the first pilot paradigm with which we can test learning performance in an automated manner. In this task zebrafish are shown animated images of a group of zebrafish on one side of the tank for a short period of time. The images disappear and remain hidden during the inter-stimulus interval after which they reappear but on the opposite side of the test tank. The experimental zebrafish are expected to learn this presentation pattern and swim to the opposite side after the disappearance of the stimulus fish. Given that both stimulus delivery (zebrafish images are presented by the computer) as well as response quantification (the location of the experimental zebrafish measured by tracking software) is computer automated, the paradigm can be run in a massively parallel manner and thus is appropriate for high throughput applications. Last, in a series of studies we have discovered significant strain differences among zebrafish populations in the way they respond to acute alcohol treatment and we are currently analyzing the effect of chronic alcohol treatment on these strains using the above tests and others. These studies also utilize computer animated images as well as computer aided automated quantification of behavior.

Given the increasingly successful application of zebrafish in behavioral neuroscience research in our and other laboratories and the feasibility of automated behavioral paradigms, we suggest that drug or mutation screening is within our reach with zebrafish and that this species will be an excellent model organism with which the pharmacology and genetics of complex brain function may be analyzed.

Approaches to analyse mood disorders in zebrafish

W.H.J. Norton[1], K. Webb[1], M. Harris[2], N. Rohner[2], C. Nüsslein-Volhard[2], J. Ninkovic[1], A. Folchert[1], and L. Bally-Cuif[1]
[1]*Department of Zebrafish Neurogenetics, Helmholtz Centre Munich, Munich, Germany, william.norton@helmholtz-muenchen.de*
[2]*Max Plank Institute for Developmental Biology, Tuebingen, Germany*

Mood disorders are debilitating diseases that show high prevalence and comorbidity in society. However, compared to other systemic diseases of the brain, the underlying genetic and neurological defects leading to mood disorders are not well understood. A lack of knowledge about their aetiology has hampered the development of drug treatments, which currently show varying efficacy and numerous side effects. Despite their poor efficiency, treatments for mood disorders still account for the largest percentage of pharmaceutical treatments prescribed, suggesting that a better understanding of the underlying pathology is crucial.

One of the aims of our research is to use zebrafish as a model organism to uncover and analyse the neural circuits that mediate mood disorder formation. The amenability of zebrafish for genetic manipulation, live imaging studies using fluorescent reporter proteins and the availability of a large mutant collection makes zebrafish an extremely appealing behavioural model. Work in our lab has already established several tests that model mood disorders in adult zebrafish, including drug addiction, boldness, and aggression. Using these protocols, we take a forward genetic approach to identify novel candidate genes which impact on mood disorder formation.

In a first set of experiments, we established a biased conditioned place preference (CPP) assay to measure drug addiction in zebrafish. Using this technique, we screened for and characterised novel zebrafish mutants that fail to change place preference following administration of the prototypical pyschostimulant amphetamine. This assay allows us to investigate the rewarding aspects of amphetamine addiction. As part of the Tuebingen screen 2005/6 we identified 6 novel mutant families that fail to become addicted to. We are in the process of further characterising one of these mutant families and are currently identifying the mutated locus by positional cloning. In parallel, we are using microarrays to examine the effect of amphetamine on the mutant and wild-type brains. A comparison of the brain transcriptome of wild-types and mutants with and without amphetamine allows us to specifically identify genes which mediate the rewarding effect of amphetamine administration. The specific aim of this set of experiments is to identify genes and pathways involved in the aetiology of addiction, with the hope of thereby identifying novel drugable targets.

In a parallel set of experiments, we have adapted an assay to measure aggression in adult zebrafish. In our assay, mirror induced stimulation is used to measure the amount of time spent attacking a perceived intruder fish in a mirror. The behaviour of single adult fish is recorded for ten minutes, and then films are replayed at a slow frame speed so that their behaviour can be analysed. Using this assay, we have identified a novel aggressive mutant, *spiegel*. *spiegel* mutants show high mortality and bite marks on the flank of conspecifics during mating. In our aggression test set-up, and in contrast to wild-type fish, *spiegel* fail to down-regulate aggression following initiation.

Using our conditioned place preference tanks, we have also measured the boldness of mutant fish. The place preference tank is divided into two separate areas: a brown side and a white side which contains two frightening black dots. Wild-type fish placed in the tank show a clear place preference, and spend approximately 85% of time on the brown side and 15% of time on the frightening white side. We reasoned that if *spiegel* is indeed bolder than wild-type, then they would modify their place preference and spend increasing amounts of time on the white side of the tank. Using this protocol we find that *spiegel* is indeed bolder than siblings.

spiegel harbours a hypomorphic mutation in *fgf receptor 1* and are both adult viable and morphologically normal. We are also dissecting the underlying neural defects leading to increased aggression through in situ hybridisation analysis. Expression of the Fgf targets Dusp6 and Phospho-ERK localises the Fgf signalling defect to a small nucleus in the inferior hypothalamus, the periventricular nucleus. This hypothalamic area has already been linked to the control of aggression in another teleost, the sunfish [1]. In parallel, the expression of the serotonin transporter gene, *serta*, is upregulated in the dorsal raphe nucleus. This suggests that a modification in 5-HT signalling underlies the behavioural changes in *spiegel*. In order to test this hypothesis, we are currently manipulating 5-HT activity within mutants, by applying the selective serotonin reuptake inhibitor Fluoxetine hydrochloride in tank water. Both wild-type and *spiegel* will be treated with a short pulse of Fluoxetine and then assayed for aggressive behaviour.

Taken together, our analysis of the *spiegel* mutant provides the first evidence that a reduction of Fgf signalling in the vertebrate brain can lead to increased aggression levels.

In summary, one of the major aims of work in our laboratory is to further develop behavioural tests for adult zebrafish. These projects will expand the number of species available for modelling the genetic and neurological changes leading to mood disorder formation in humans, and potentially identify new drug targets for their treatment.

References

1. Demski, L.S., Knigge, K.M. (1971). The telencephalon and hypothalamus of the bluegill (Lepomis macrochirus): evoked feeding, aggressive and reproductive behavior with representative frontal sections. *J Comp Neurol* **143**, 1-16.

The home cage as starting point for innovative concepts in behavioural phenotyping Symposium

Leonie de Visser
Dept. Animals, Science and Society, Utrecht University, Utrecht, The Netherlands, leonied@gmail.com

Minimal human intervention, handling and transport; undisturbed and continuous behavioural recordings; measurements of long-term and circadian processes. These advantages have made the home cage increasingly popular and acknowledged as testing environment for behavioural phenotyping purposes. Flanked by an upsurge in the number of automated home cage recording systems available on the market, the home cage is more often becoming implemented in behavioural assays and test batteries. Measuring activity under baseline conditions in the home cage facilitates the interpretation of outcomes from other, novel environment tests. However, the home cage can be equipped to serve as a testing environment for an extended number of behavioural domains, such as anxiety and cognition. This symposium presents the most recent innovations and applications in the field of home cage studies. The program is aimed at addressing different types of automated home cage systems that each have their specific methodology suitable for answering the specific questions of their users. Moreover, the symposium will focus on a variety of behavioural domains, such as locomotor activity, emotional learning, operant conditioning and anxiety. The assessment of home cage behaviour presented here is implemented in neurobiological studies combined with genetic strategies and specific disease-related research.

Symposium contents

Fully Automated 24/7 Behavioral Screening for Mutations in Targeted Cognitive Mechanisms in the Mouse

C.R. Gallistel, A.P. King, A.M. Daniel, E.B. Papachristos, and F. Balci
Department of Psychology & Center for Cognitive Science, Rutgers University, New Brunswick, NJ, USA,
galliste@ruccs.rutgers.edu

Current screens for memory defects in mutant mice, such as the Morris water maze, generally proceed on the assumption that memory impairment will be manifest in an altered rate of learning. The rate of learning in mutant and wild-type strains is estimated by plotting a group average measure of performance--for example, the mean latency to find the platform--as a function of trials. A problem with this approach is that the gradual approach to asymptote seen in group-average plots is an artifact of the averaging [1-3]. Thus, the measured quantity, the learning rate, does not reflect a meaningful quantity within the individual subjects. A second problem is that the trial-by-trial handling of the mice is stressful to them and wasteful of the experimenter's time. The results strongly reflect response to handling stress, and the paradigms cannot be scaled up to allow for large scale screening.

We believe there is a large latent demand for the large-scale behavioral screening of mutant mice for heritable malfunctions in the mechanisms of cognition. Seymour Benzer and his students, in their seminal use of genetics to get at the molecular biology of the circadian clock [4] have shown the power of this approach to take us from the behavioral to the molecular level of analysis.

We believe the keys to a successful screening program are: 1) the targeting of behaviorally well defined mechanisms, like the circadian clock, for which one can make physiologically meaningful quantitative measurements at the behavioral level (e.g., the measurement of the free-running period, or of the spectral sensitivity of its entrainment mechanism [5, 6].); 2) the development of automated procedures that eliminate handling of the mice during the period when behavioral measurements are made and give as many measurements as possible in as little time as possible.

Our research targets the interval timing mechanism, whose behavioral investigation was pioneered by Gibbon and Church [7, 8], and the mechanisms for estimating probabilities (relative frequencies) and the proportions obtaining between them. The physiologically meaningful quantities that we measure are the accuracy and precision of the individual subject's representation of these objective quantities (duration and relative frequency and proportion). We have developed

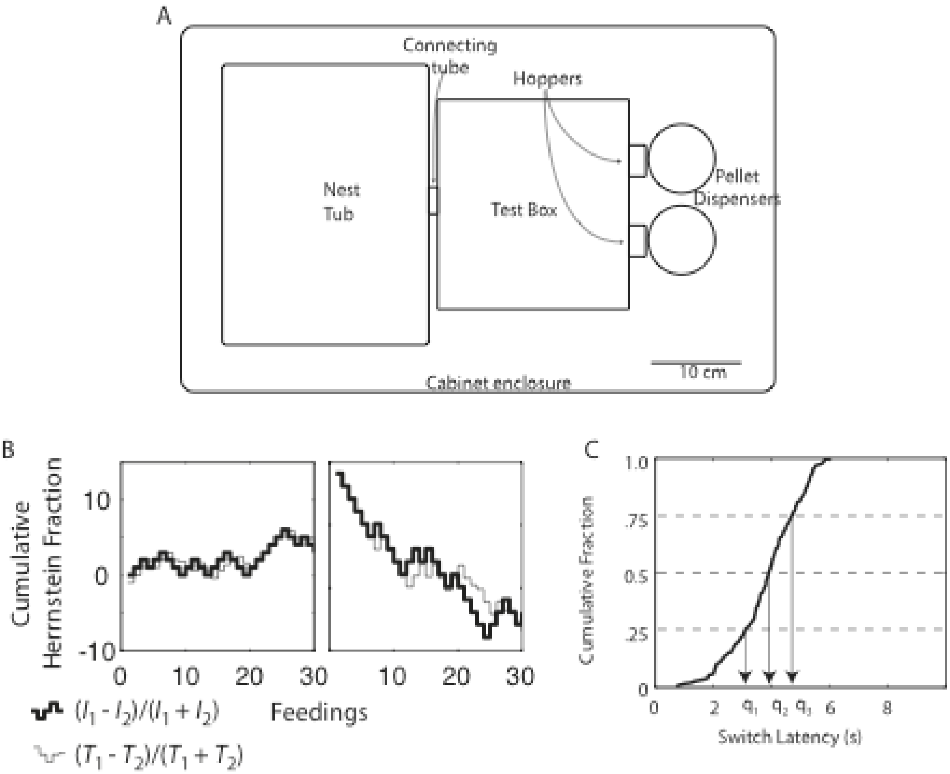


Figure 1A. *Plan of live-in test environment. A nest tub communicates with a Med Associates™ Mouse Test Box by way of a connecting tube. Test box has two illuminable feeding hoppers monitored by IR beams.* ***B.*** *Cumulative records of the Herrnstein fractions over the first 30 feedings of two experimentally naïve CB57/B6 female mice (from Gallistel et al). The Herrnstein income fraction (plotted with heavy lines) is the proportion of total pellets obtained from Hopper 1; the time fraction (light lines) is the proportion of total hopper visiting time devoted to visiting Hopper 1. When the slopes of these two cumulative records are the same, the subject is matching its visit proportion to its income proportion. (From 9)* ***C.*** *Cumulative distribution of switch latencies from a female CB57/B6 mouse, with short and long delays of 2 and 6 seconds and equal relative frequencies. The median latency (q_2) measures how accurately the subject estimates the midpoint between the delays; the inter-quartile interval ($q_3 – q_1$) measures the precision of this estimate (its variability). Balci [11] showed that the median shifts in accord with the relative frequencies of the two delays (their probability). (From 10).*

paradigms for measuring these quantities rapidly in a live-in environment, which eliminates the handling of the mice (Figure 1A). By automating every aspect of the situation, including much of the data analysis, which is conducted in quasi real time, we make it possible to do large scale screening with an equipment investment no larger than is required for many major molecular and neurobiological experimental programs.

We use the matching paradigm to measure the accuracy with which the mouse estimates the average intervals between randomly scheduled pellet releases into two different hoppers and the accuracy with which it represents the proportion between these average intervals. In the matching paradigm, the mouse adjusts the expected durations of its visits to the two hoppers so that their ratio (the proportion between the two expectations) matches the ratio of the expected intervals between pellet releases. Mice reliably exhibit matching within the first few hours in a new test environment (Figure 1B), a period during which they may remain so wary of the new environment that they eat only a few of the pellets they obtain by poking into the feeding hoppers [9].

We use the "switch" paradigm [10] to measure the accuracy and precision with which the mouse represents durations and the accuracy with which it represents a probability (relative frequency). In this paradigm, a trial begins with the illumination of the two hoppers. With some relative frequency, the trial terminates with the delivery of a pellet to, say, the left hopper after a fixed delay of, say, 2 s. With the complementary relative frequency, it terminates with the delivery of a pellet to the other hopper after a fixed delay that is longer by some fixed factor (typically in the range 1.5 to 3). The mice soon learn to begin every trial by poking repeatedly into the short-delay hopper and to switch to the long-delay hopper on those trials (long trials) when the short delay expires without the release of a pellet. The accuracy of the mouse's representation of the delays is indicated by the median of the distribution of these switch latencies; its precision by their inter-quartile interval (Figure 1C). The median switch latency also depends systematically on the relative frequency of the short and long-delay trials. It shifts toward or away from the short delay according as it is more or less probable [11].

References

1. Morris, R.W., Bouton, M.E. (2006) Effect of unconditioned utimulus magnitude on the emergence of conditioned responding. *Journal of Experimental Psychology: Animal Behavior Processes* **32**: 371-385.
2. Gallistel, C.R., Balsam, P.D., Fairhurst, S. (2004) The learning curve: Implications of a quantitative analysis. *Proceedings of the National Academy of Sciences (USA)* **101**: 13124-13131.
3. Papachristos, E.B., Gallistel, C.R. (2006) Autoshaped head poking in the mouse: A quantitative analysis of the learning curve. *Journal of the Experimental Analysis of Behavior* **85**: 293-308.
4. Weiner, J., (1999) Time, love, memory: A great biologist and his quest for the origins of behavior., New York: Alfred A. Knopf. 300.
5. Foster, R.G., (1993) Photoreceptors and circadian systems **2**: 34-39.
6. Takahashi, J.S., et al. (1984) Spectral sensitivity of a novel photoreceptive system mediating entrainment of mammalian circadian rhythms. *Nature* **308**: 186-188.
7. Gibbon, J., Church, R.M., Meck, W.H. (1984) Scalar timing in memory. In *Timing and time perception*, J. Gibbon and L. Allan, Editors., New York Academy of Sciences: New York. p. 52-77.
8. Church, R.M., Broadbent, H.A., Gibbon, J. (1992) Biological and psychological description of an internal clock. in *Learning and memory: The behavioral and biological substrates*, I. Gormezano and E. Wasserman, Editors., Erlbaum: Hillsdale, NJ. p. 105-128.
9. Gallistel, C.R., et al. (2007) Is matching innate? *Journal of the Experimental Analysis of Behavior* **87**: 161-199.
10. Balci, F., et al. (in press, August 2007) Interval timing in genetically modified mice: a simple paradigm. *Genes, Brains and Behavior*.
11. Balci, F. (2007) *Risk assessment in man and mouse*. PhD Thesis. Rutgers University, New Brunswick

Learning (in) the PhenoTyper: an integrative approach to conducting cognitive behavioural challenges in a home cage environment

R.C. de Heer, M.Schenke, W.W. Kuurman, and B.M. Spruijt
Department of Animals, Science and Society, Utrecht University, Utrecht, The Netherlands, r.c.deheer@uu.nl

Mice are extensively used in neurobehavioural research as a model for many human (behaviour) disorders. In order to characterize the behaviour of certain mouse strains, many uni-dimensional tests have been developed, which allow characterizing the mouse on a few aspects of the behavioural domain of interest. In addition Stand alone tests have been combined into test batteries. Mice will then be sequentially exposed to different tests and are usually handled and "transported" to a different experimental set-up. But the validity of studies of animal behaviour ignoring the interaction of more motivational systems, either by a strong focus in the ethogram or a focus in the design on predominantly activity, is limited. An alternative approach is the modification of existing "traditional" tests into more complex tests by trying to incorporate more behavioural domains into it. The discriminative elevated plus-maze and the discriminative avoidance task using a Y-maze are examples of such modifications.

e developed a reliable tool for observing and analysing various behavioural systems in a home-cage-like environment, without any handling and/or transport of animals. This High Throughput Phenotyping-system (HTP-system) is an extension of the PhenoTyper-system (Noldus, Wageningen, The Netherlands).

In addition to the measurement of activity and anxiety we have designed and implemented two fully-automated cognitive challenges for mice in this set up. Both tests have in common that we do not employ techniques to enhance the motivation of a mouse, e.g. food deprivation or other strong aversive stimuli. In fact the cognitive performance relies on the willingness of the animal to collaborate The tests will commence automatically at a certain, predefined time (e.g. several days) in the PhenoTyper, which makes it possible to compare behaviour of mice during the test with that of base line behaviour before and behaviour after the test. Furthermore, the HTP-system allows the mouse to display more of its natural behavioural repertoire and thus enables the researcher to carry out more complex analyses and allows studying the potential interaction of different behavioural dimensions with data from the same experiment.

The first task is a discriminative avoidance task, in which a mild aversive stimulus is used. After 4 days of continuous monitoring, the number of shelter entries using both the left and right entrance was determined; the entrance that was used most frequently was considered to be the preferred entrance and was subsequently designated as incorrect on day 5 and day 6. If the mouse used this incorrect entrance on day 5 and day 6, a bright light inside the shelter was switched on. The light stayed on as long as the mouse remained inside the shelter after making the incorrect choice. Using the non-preferred entrance had no consequences. In front of the shelter we defined two areas and each area is linked to either the left or the right shelter entrance. Whenever the mouse enters the zone linked to the preferred and hence "incorrect" entrance, a cue (a short light flash) was given. This test allows us to analyse discrimination and avoidance learning, in interaction with anxiety and locomotion.

The second test we designed is a free-operant conditioning task, consisting of several sessions in which the mouse has to learn that jumping onto its shelter will result in dropping a sucrose pellet elsewhere in the cage (an area around a pellet dispenser adjacent to the feeder. The mouse first has to go to the area where the pellet is dropped, before the next trial could be initiated. Prior to testing the mice are not food deprived and during the test mice still had access to food ad libitum. This experiment demonstrates place preference in interaction with activity, locomotion and motivation for the food incentive.

We will present data of two inbred strains of mice (C57BL/6 and DBA/2), showing that it is not only possible to carry out fully automated behavioural challenges in a homecare-like environment in relatively short period of time, but that is possible to separate these strains from each other based on their performance in these tasks.

Differential involvement of the central amygdala in appetitive versus aversive learning in mice trained in the IntelliCage system

E. Knapska[1], F. Neuhäusser-Wespy[3], H.-P. Lipp[3], L. Kaczmarek[2], and T.Werka[1]

[1]Department of Neurophysiology and [2]Department of Molecular and Cellular Neurobiology, Nencki Institute, Warsaw, Poland; [3]Division of Neuroanatomy and Behavior, Institute of Anatomy, University of Zurich,Zurich, Switzerland

The IntelliCage is a large home cage containing four operant conditioning chambers placed in the corners. Access into each chamber is provided via a tubular antenna reading transponder codes. Before entering the cage every mouse is injected subcutaneously with a glass-covered microtransponder. This passive transponder emits a unique animal identification code when activated by a magnetic field. The design of the cage restricts access to the learning chamber for a single mouse only. Each chamber contains two openings that give access to the nipples of drinking bottles. These openings are crossed by photobeams recording nose-pokes of the mice. Access to the tubes can be barred by small motorized doors. Moreover, an aversive stimulation can be delivered in forms of air-puffs directed to the head of the mouse through tubing controlled by electric valves. The whole set-up of the IntelliCages is controlled by a microcomputer recognizing visits, nose-pokes, and tube-lickings of individual mice, and delivering reward (by opening the access to water after a nose-poke) or punishment (by applying air puffs after a mouse enters the test chamber) according to preprogrammed schedules depending on the assignment of the mice to different test groups within the same cage. Activity and learning of up to 16 transponder-tagged mice per cage can be continuously monitored and controlled by a computer without human interference. Thus, the system allows automated cognitive and behavioral screening of mice living in social groups.

I would like to present one of many possible applications of the IntelliCage system. The aim of this study was to compare the engagement of various nuclei of the amygdala in appetitive and aversive instrumental training procedures [1]. We used the IntelliCage system to balance the conditions of those trainings. The mice were exposed to a place preference and place avoidance training. In the place preference training, the animals were supposed to associate the sweetened water with a specific corner within the cage (appetitive motivation), whereas in the aversive training they were learning to avoid a corner where the air-puffs were applied (aversive motivation). Appetitive and aversive training was performed in three cages each. Two 8-day sessions were carried out: (i) the first one with the place preference group and their controls and (ii) the second one with the place avoidance group and their controls. Every cage contained four or five mice assigned to the different treatment groups. A session started with a 48-h adaptation period to the cage during which the mice learned to open the gates barring access to plain water from both openings by means of nose-pokes. For the next 3 days, the animals were temporarily deprived of water. They had access to water for 2 h per day, at the same time of the active phase every day. This procedure evoked intense consummatory activity during a limited time span. On the following day at the same time, the mice were assigned to different treatment groups within the cages. One group was designed as the "place preference group". They received sweetened water (10% sucrose) in the corner least preferred during the previous drinking sessions. In the "place avoidance group", the mice received an air puff when entering the corner that was the most preferred during the previous drinking session. Control groups consisted of mice that were in the same cages as the experimental mice, but obtained sweetened water in all four corners or received no air-puffs but only plain water, respectively. The whole experimental schedule in the second and third cage for each type of the training was shifted by 1 and 2 h, respectively, with respect to the schedule in the first cages in order to allow timely removal of animals for immunohistochemistry. All animals learned the required tasks very effectively.

Then, these behavioral paradigms were used to map the patterns of c-Fos expression in the amygdala. The brains for immunocytochemistry were taken on the first day of place conditioning, from three mice from each cage. We have found much more intense c-Fos expression in the medial part of the central amygdala after the appetitive training as compared to the aversive training. In contrast, the similar level of c-Fos expression was evoked by both types of training within the lateral nucleus of the amygdala. The data support the hypothesis that the central nucleus of the amygdala is particularly involved in appetitively motivated learning processes.

Supp. by FP6 037965 and NCCR "Neural Plasticity and Repair"

References

1. Knapska, E., Walasek, G., Nikolaev, E., Neuhäusser-Wespy, F., Lipp, H.P., Kaczmarek, L., Werka, T. (2006). Differential involvement of the central amygdala in appetitive versus aversive learning. *Learning & Memory* **13(2)**:192-200.

Genetic dissection of motor activity and anxiety-related behaviors in mice using an automated home cage task

Martien J.H. Kas[1], Berend Olivier[2], and Annetrude (J.G.) de Mooij-van Malsen[1]
[1]Department of Neuroscience and Pharmacology, Rudolf Magnus Institute of Neuroscience, UMC Utrecht, The Netherlands.
[2]Department of Psychopharmacology, Utrecht Institute for Pharmaceutical Sciences and Rudolf Magnus Institute for Neurosciences, Faculty of Pharmaceutical Sciences, Utrecht University, Utrecht, The Netherlands,
m.j.h.kas@umcutrecht.nl

Introduction

Family and twin studies have revealed that genetic factors play a major role in psychiatric disorders, however, attempts to find susceptibility genes for these complex disorders have been largely unsuccessful. Therefore, new research strategies are required to tackle the complex interactions of genes, developmental, and environmental events. Recently, we have proposed a behavioral domain concept that focuses on the genetics of behavioral domains relevant to both animal behavior and across human psychiatric disorders [1]. We believe that interspecies trait genetics rather than complex syndrome genetics will optimize genotype–phenotype relationships for psychiatric disorders and facilitate the identification of biological substrates underlying these disorders. The development of automated paradigms that address these behavioral domains is a crucial step in this translational research field.

The way forward

Identification of novel genetic loci in animal models for neurobehavioral traits relevant to psychiatric disorders relies on the fact that these traits are truly translatable across species. Once found, one can apply sensitive genetic strategies to these traits in order to unravel the underlying mechanisms. With the availability of a large variety of inbred mouse strains and their recently known genome sequences, mouse genetics offer a challenging way to study complex behavioural traits. For example, in contrast to patient populations, mouse strains can be used to control for phenotypic and genetic heterogeneity as well as for complex gene-environment interactions. Interestingly, recent studies have shown that genetic variation associated with psychiatric disorders affect analogous neural circuits and behavioural traits in mice and men, demonstrating that mouse models can contribute to systematic searches for genetic determinants of psychiatric disorders [2].

In general, rodent species have an innate preference for sheltered places that have lower light intensities than the outside-world and that provide a sense of safety via body contact with the shelter area surface (thigmotaxis). The assessment of this anxiety-related behavior is generally measured in relative short-lasting laboratory tasks and is highly dependent on strain differences in motor activity levels. In light of this, we have recently designed an automated home cage environment to assess separate behavioral domains over the 24-h day [3, 4, 5]. For instance, a hungry organism searching for food depends on an efficient exploration strategy in which finding the food resource in an appropriate period of time needs to be balanced against the risk of being exposed to potentially threats, such as predators. Thus, exploration for food relies on a balance between movement throughout the environment and avoidance behavior. To assess these behavioral domains as a function of time of day, a home cage environment for mice was designed with a sheltered and a non-sheltered feeding platform that would allow dissociation of the preference for shelter during feeding and for motor activity levels over several days and with minimal human disturbance [5].

For mouse genetics of these behavioral domains, there is an increasing appreciation of the properties of the set of mouse inbred strains which have been established over the last century of mouse genetics. Data are accumulating on each of this diverse collection of over 500 strains allowing strains to be chosen that cover a range of phenotypic variation in whatever phenotype is of interest. Traditionally, such strain combinations would be used to set up a cross or segregating population for genetic mapping purposes. More recently, Genetic Reference Populations (GRPs) with more optimal genetic properties are available or under construction. The prototype is the Recombinant Inbred (RI) panel which is generated from a cross between two inbred strains followed by an F1 intercross and 20 generations of inbreeding. The best characterized mouse RI panel, derived from C57BL/6J and DBA/2J (BXD) strains, has been a workhorse of behaviour genetics since the early 1990s. The BXD panel, of (until recently) 35 lines, gives only a coarse genetic resolution. However, several aspects of this picture have recently changed. One is the idea of treating transcript abundance and protein abundance or modification as phenotypes in their own right. This kind of genetical-genomics is an extremely promising way of examining networks of function. Although it is possible with conventional genetic crosses or outbred populations, using (effectively immortal) inbred GRPs allows much more value to be extracted from each data set. This is also true of other phenotypic data, and there is a renewed interest in the importance of accumulating data from many investigators. It is now possible to not only genetically map Quantitative Trait Loci (QTLs) on the BXD RI panel (recently expanded to 80 lines), but also to correlate new data with a large database of phenotypic data including gene expression data on several tissues. This is a major aid to positional cloning projects and multivariate analysis approaches currently being explored offer a way to assign some idea of function to many of the genes whose function is currently unknown. Other genetic reference populations have been developed. For example, chromosome substitution (consomic) panels have been generated for two strain combinations. This is done by repeated backcrossing to produce strains each with a single chromosome of one strain on the background of another. This is attractively simple to analyze and offers the simplification of multi-locus traits with only a single chromosome segregating [6].

Results and conclusion

By testing a panel of 21 chromosome substitution strains in a wide variety of traditional and in an automated home cage environment, we have shown that behavioral components can be genetically dissociated. For example, we have shown that motor activity levels are under different genetic control than the preference to shelter by using a novel automated home cage task [5]. Further genetic mapping of these chromosomal regions revealed genetic loci that are syntenic with human linkage regions for mood disorders. Candidate gene selection within the QTL-intervals can nowadays be facilitated by combining quantitative phenotypic data from inbred mouse strains with their online available genome sequences. Single Nucleotide Polymorphism (SNP) databases provides gene-by-gene, SNP-by-SNP distribution patterns for various inbred lines and allow pinpointing SNP's in the QTL-regions that are associated with the behavioral trait of interest. Furthermore, biological pathway analysis can further be applied to provide additional candidate gene information. Subsequently, homologous candidate genes can then be tested in DNA samples from well-characterized psychiatric patient populations. In this way, interspecies genetics offers a great opportunity to translate essential behavioral traits in animals to human psychiatric disorders and to further understand the mechanisms underlying these traits.

References

1. Kas M.J., Fernandes C., Schalkwyk L.C., and Collier D.A., (2007) Genetics of behavioural domains across the neuropsychiatric spectrum; of mice and men. *Molecular Psychiatry,* **12**(4), 324-30.
2. Chen Z.Y., Jing D., Bath K.G., Ieraci A., Khan T., Siao C.J., Herrera D.G., Toth M., Yang C., McEwen B.S., Hempstead B.L., Lee F.S., (2006) Genetic variant BDNF (Val66Met) polymorphism alters anxiety-related behaviour. *Science*, **314**(5796), 140-3.
3. Kas MJ, Van Ree JM. Dissecting complex behaviours in the post-genomic era. (2004) *Trends Neurosci.* **27**(7):366-9.
4. de Visser L, van den Bos R, Kuurman WW, Kas MJ, Spruijt BM. (2006) Novel approach to the behavioural characterization of inbred mice: automated home cage observations. *Genes Brain Behav*. **5**(6):458-66.
5. Kas M.J., de Mooij-van Malsen J.G., Olivier B., Spruijt B.M., Van Ree J.M. "Differential genetic regulation of motor activity and anxiety-related behaviors in mice using an automated home cage task. *Behavioral Neuroscience*, in press
6. Singer J.B., Hill A.E., Burrage L.C., Olszens K.R., Song J., Justice M., O'Brien W.E., Conti D.V., Witte J.S., Lander E.S., Nadeau J.H., (2004) Genetic dissection of complex traits with chromosome substitution strains of mice. *Science*. **304**(5669), 445-8.

Home cage testing of impulsivity

S. Koot[1,2], W. Adriani[1], R. van den Bos[2], and G. Laviola[1]
[1]*Behavioural Neuroscience Section, Dept. Cell Biology & Neurosciences, Istituto Superiore di Sanità, Roma, Italy, walter.adriani@iss.it*
[2]*Ethology and Welfare, Faculty of Veterinary Medicine, Utrecht University, The Netherlands*

Lack of self-control or impulsive decision making may be an important symptom of psychiatric disorders, such as Attention Deficit Hyperactivity Disorder [2,7]. Animal models are crucial in studying the underlying neurobiology. In the intolerance-to-delay (ID) task [3,5], subjects may choose between a late-large and a soon-small reward. Impulsive subjects are intolerant to the forced waiting for the large reward [4,5]. The rats' performance on the ID-task is investigated by placing the animals in individual operant chambers for a short period daily [4,5]. However, stress caused by handling and by novelty might influence their performance. Therefore, a new computer-controlled operant panel was developed, which can be placed inside the home cage, enabling the rat to operate it 24 hours/day. Here we report results of a pilot experiment using this panel in an ID-protocol.

Materials & methods

Subjects

Four adult male rats (Harlan, Italy; mean weight 429gr) were kept in an air-conditioned room (temperature 21 ± 1°C, relative humidity 60 ± 10%), on a 12-hr reversed light-dark cycle (lights on at 8.00 pm). Prior to the experiments animals were housed in pairs, but from the start of the protocol animals were singly housed. Water was available ad libitum, whereas food (Altromin-R, A. Rieper S.p.A., Vandoies, Italy) was available ad libitum until the start of the protocol. Rats had previous experience in impulsivity tasks in a classical skinnerbox setting, two months prior to the present pilot.

Apparatus

The testing apparatus consisted of one computer-controlled operant panel for each of the subjects, placed in a Macrolon III cage with sawdust bedding. The panel contains two nose-poking holes, hole lights, a chamber light, a feeder device, a food-magazine where pellets (#F0021-J Dustless Precision Pellet 45 mg, BioServe, Frenchtown, USA) are delivered, and a magazine light. The panel was attached through an interface to a PC, where a custom-made software controlled and recorded all events. Nose-poking in one of the two holes of the panel resulted in the delivery of five pellets (large reward), whereas nose-poking in the other hole resulted in the delivery of one pellet (small reward). After nose-poking and before food delivery, the hole light was turned on for 1s. Following food delivery the magazine light was turned on for 90s, during which nose-poking was recorded, but was without scheduled consequences (timeout). The magazine light was then turned off, the chamber light was turned on, and the system was ready for the next trial.

Protocol impulsivity test

On day 1, animals were placed in the cages containing the panel, which occupied one fourth of the total living area. The adaptation period started with 24 hours of access to ad libitum regular food pellets (Altromin-R) and BioServe pellets from the panel. On day 2, Altromin-R was removed for 24 hours, while animals still had access to BioServe pellets. Then, 12 hours of food deprivation followed in order to increase their motivation to work for food delivery.

During the subsequent training and testing phases, animals had only access to BioServe pellets during the sessions, by operating the panel, and to a limited amount of Altromin-R after each session: two 1h sessions were run daily between 9.00-10.00 and 18.00-19.00 [for arguments see ref. 6]. After each session, the total intake of BioServe pellets was calculated per individual, and additional food was given to meet their daily nutritional needs (details available on request). The end of a session was indicated by switching off all panel lights plus the delivery of the additional Altromin-R pellets. Training lasted until all subjects reached a significant preference for the large reward.

During the testing phase, a signalled delay was added to the 1s-interval, normally scheduled between nose-poking and large-reward delivery. The hole light was kept on during the entire length of this delay. The small reward delivery was unchanged. Hence, animals had a choice between a "large & late" (LL) and a "small & soon" (SS) reward. The delay length was fixed for daily sessions and was changed over days: 15s on the first day, followed by delays of 45s, 75s, 105s and 150s on subsequent days.

Results

Following four training sessions, all rats showed a significant preference for the large over the small reward (average choice of 94.4 ± 5.3% for the large reward). This finding replicated previous experiments in our lab [1] and also indicates that animals remain to probe the outcome of nose-poking at the other hole.

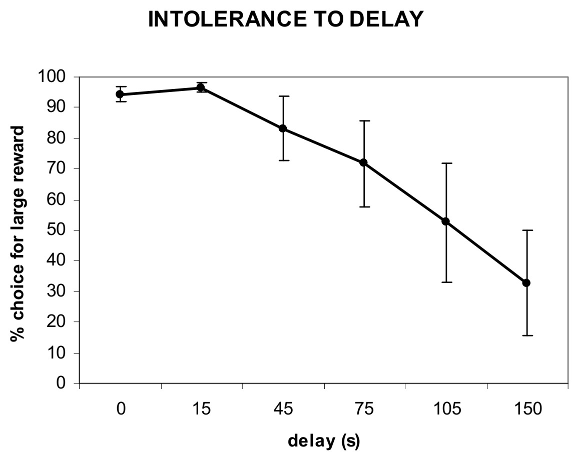


Figure 1. *Choice behaviour in rats (n = 4) tested with the intolerance-to-delay (ID) protocol, shown during daily sessions in the home cage situation. Data represent the mean (±SEM) choice (%) for the larger reward per day, delivered after a delay.*

When delays were gradually increased over days (Figure 1), rats showed a shift towards more SS choices at the longest delays [cf. 4,5]. The protocol lasted 9 days in total.

Discussion

The present pilot experiment shows that it is in principle possible to measure impulsive behaviour in a home cage setting. Future experiments are directed at validating this approach.

References

1. Adriani, W., Rea, M., Baviera, M., Invernizzi, W., Carli, M., Ghirardi, O., Caprioli, A., Laviola, G (2004). Acetyl-L-carnitine reduces impulsive behaviour in adolescent rats. *Psychopharmacology* **176,** 296-304.
2. Bechara, A. (2004). The role of emotion in decision-making: evidence from neurological patients with orbitofrontal damage. *Brain and Cognition* **55**, 30-40
3. Evenden, J.L. (1999). Varieties of impulsivity. *Psychopharmacology* **146,** 348-361.
4. Evenden, J.L., Ryan, C.N. (1996). The pharmacology of impulsive behaviour in rats: the effects of drugs on response choice with varying delays of reinforcement. *Psychopharmacology* **128,** 161-170.
5. Evenden, J.L., Ryan, C.N. (1999). The pharmacology of impulsive behaviour in rats VI: the effects of ethanol and selective serotonergic drugs on response choice with varying delays of reinforcement. *Psychopharmacology* **146,** 413-421.
6. Strubbe, J.H., Woods, S.C. (2004). The timing of meals. *Psychological review* **111,** 128-141.
7. Swanson, J., Castellanos, F.X., Murias, M. (1998). Cognitive neuroscience of attention deficit hyperactivity disorder and hyperkinetic disorder. *Current Opinion in Neurobiology* **8**, 263-271.

The PhenoTyper automated home cage environment as a high throughput tool to detect behavioral abnormalities in mutant mice

M. Loos
Center for Neurogenomics and Cognitive Research, VU University, Amsterdam, The Netherlands, maarten.loos@falw.vu.nl
The Neuro-Bsik Mouse Phenomics consortium[1]

Over the last decades, the high throughput toolbox of neurobiologists to measure molecular changes in the brain has grown rapidly. The molecular effects of induced genetic mutations or pharmacological treatment relevant to psychiatric disorders can now be studied on a large scale. In contrast, to investigate the behavioral consequences of genetic mutations or pharmacological treatment, researchers rely on multiple stand-alone tests, each of which taxes a specific aspect of a different behavioral domain. It is clear that these conventional stand-alone tests give highly valuable detailed information about specific psychological constructs. However, for the purpose of phenotyping larger numbers of novel mutants or novel pharmaceuticals, performing a battery of these conventional tests has two major drawbacks. First of all, performing a battery of tests requires a substantial amount of time and resources. Secondly, it has proven to be difficult to reproducibly measure behavioral effects in different laboratories, especially subtle ones [1]. This difficulty in reproducibility may partly be due to a multitude of human-animal interactions required for these behavioral tests. Therefore, automation may be a way of reducing the required amount of time and resources involved with behavioral phenotyping. In addition, automation will circumvent human-animal interactions, and thereby holds the promise to increase reproducibility of behavioral results across laboratories. The development of automated phenotyping strategies has only started recently, for instance using automated home cage observations[2]. More research is needed, especially to increase throughput and to investigate the sensitivity and reproducibility of these automated behavioral tests.

Within the framework of the Dutch Neuro-Bsik Mouse Phenomics project, an automated high throughput screening protocol was developed, consisting of an automated home cage environment (PhenoTyper®) and separate home cages equipped with running wheels. Neuro-Bsik Mouse Phenomics is a Dutch consortium of 11 academic research groups and two companies. The consortium aims to contribute to the understanding of brain disorders by developing novel mouse models for brain disorders. These models will be initially identified using the automated high throughput screening protocol to analyze mouse behavior and subsequently be analyzed in depth at different levels of complexity using the specific expertise of each of the academic partners.

For the aims of the consortium, it is essential to investigate whether the developed automated high throughput screening protocol can discriminate mouse behavior with sufficient resolution, e.g. mutant from wild type mice or specific behavioral differences between common inbred lines of mice. Secondly, if differences in behavior between genotypes are detected, the high throughput screen should indicate in which behavioral domain (e.g. circadian rhythm, anxiety, learning) further in-depth analysis might be fruitful.

To this end, the behavior of around fifty genetically diverse mouse lines is currently being measured using the automated high throughput screening protocol comprising engineered mutants, eight common inbred strains and the panel of BXD recombinant inbred strains. After acclimatization to the facility for at least one week, each mouse is individually housed in one of the forty-eight PhenoTyper cages for six and a half days. Using newly developed soft- and hard-ware, i.e. a video camera, Lick-O-meters and visual and acoustical stimuli, numerous measures of spontaneous behavior are recorded throughout these six days. During the last three dark phases, mice are exposed to three tests tax aspects of cognition and anxiety; an instrumental conditioning task using a device that dispenses palatable food, an avoidance learning task using illumination in the shelter, and an anxiety test using a bright spot of light on the cage floor. To automate these tasks, a series of new software applications have been developed for high-throughput data acquisition, data processing and statistical analysis. A ‘dashboard’ program enables the user to remotely monitor the mice in the PhenoTyper cages and the progress of all trials, and to perform remote diagnostics of the system. After the 6 days in the PhenoTyper, mice are transferred to new home cages containing a running wheel. For the first 5 days, mice are left to habituate to these cages under a standard light-dark regime. During the final 5 days lights remain off and the free running rhythm of mice is assessed.

Ultimately, in an automated process, the behavioral measures of each individual mouse are aggregated into a phenogram, to allow comparison with a reference mouse (i.e. a wild type strain).

The behavioral measures currently being obtained in the high throughput screening protocol will be compared to known behavioral abnormalities of these strains, as measured with several established stand-alone tests. During the conference presentation I will address whether the current automated high throughput screening protocol allows reliable discrimination of common inbred lines and whether some behavioral measures can be validated by the standard stand alone testing protocols.

References

1. Crabbe, J.C., Wahlsten, D., Dudek, B.C., 1999. Genetics of mouse behavior: interactions with laboratory environment. Science (New York, N.Y 284, 1670-1672.
2. de Visser, L., van den Bos, R., Kuurman, W.W., Kas, M.J., Spruijt, B.M., 2006. Novel approach to the behavioural characterization of inbred mice: automated home cage observations. Genes Brain Behav 5, 458-466.

[1]The Neuro-Bsik Mouse Phenomics consortium consists of the following academic and commercial partners: Erasmus University Medical Center, represented by the departments of Neuroscience (Prof. Dr. C.I. de Zeeuw, Dr. Y. Elgersma, Prof. Dr. J.G.G. Borst) and Cell Biology and Genetics (Dr. ir. N.J. Galjart, Prof. Dr. G.T.J. van der Horst). Vrije Universiteit Amsterdam, represented by the departments of Functional Genomics (Prof. Dr. M. Verhage), Experimental Neurophysiology (Prof. Dr. A.B. Brussaard) and Molecular and Cellular Neurobiology (Prof. Dr. A.B. Smit). Utrecht University, represented by the department of Animals, Science and Society (Prof. Dr. B.M. Spruijt). University of Amsterdam, represented by the Swammerdam Institute for Life Sciences (Prof. Dr. C.M.A. Pennartz). Netherlands Institute for Neuroscience (Dr. C.N. Levelt). Synaptologics BV. Noldus Information Technology.

Experience Labs: towards a contextualized understanding and evaluation of user experiences Symposium

P. Markopoulos
Dpt. Of Industrial Design, Eindhoven University of Technology, The Netherlands, p.markopoulos@tue.nl

Ambient Intelligence and related visions such as pervasive computing, ubiquitous computing, etc., signal a move towards embedding information and computational technology in our social and physical interactions making it an inseparable part of our daily life. For researchers, designers or technologists attempting to design, analyze, engineer or create user experiences this transition poses serious methodological challenges. These are discussed briefly below:

- Designers or researchers are expected to study and analyze a situation that does not yet exist. Currently only modest scale demonstrators of Ambient Intelligence technologies have been created rather than realistic deployments. Test users can only reflect upon their understanding of designed experiences by extrapolating from brief encounters with experimental technologies and with partial representations of this future technological domain divorced from the contexts of their daily lives. User studies conducted within such constraints lack external validity. More critically, they cannot produce dense explanations of the phenomena surrounding the user experience and are too limited to further our current understanding of these experiences and to drive related design efforts.
- Existing evaluation methods and methodological research have focused on task-oriented interaction, usually embedded within a short time span. Extending characterizations and evaluation methods to address user experiences as they occur in context, reflecting social interactions between several participants, requires scaling up the sampling of data and the richness of the data collected through user studies.
- The problematic of understanding the user experience extends beyond usability and human factors accounts. Concepts (overlapping) such as persuasion, fun, enjoyment, engagement, flow, trust, are not yet sufficiently understood; presenting thus a vibrant field of research in defining, operationalizing and measuring related concepts.
- The eventual form of the interactive experience will depend as much on any particular interactive product as upon the technological, social and the physical contexts in which this product will be experienced. Studies of the user experience must be able to account for, capture and investigate this variability allowing experimenters to manipulate and control those environments or, when working in the field, to capture sufficient contextual information about it.

A growing number of research initiatives attempts to address these challenges. First, considerable work is going into understanding the concept of user experience, providing theoretical accounts for the range driving the development of new technologies, like fun, connectedness, engagement, etc, involving survey measures but also using physiological parameters of humans (e.g., galvanic skin conductivity, heart-rate, etc.) as ways to assess aspects of those experiences.

Laboratory infrastructures have been created allowing long term experimental deployments of ambient intelligence technologies. A leading example of such an infrastructure is the Future Home at Georgia Tech.

Other laboratories focus on providing a realistic simulation of a target environment, be that a home, an office or a hospital environment. Inside such simulated labs, often having multiple areas, it is often possible to modify this environment for the purpose of an experiment (e.g., modifying the layout) and to manipulate several environmental variables, e.g., lighting, temperature. A pioneering example of such an infrastructure is the Home Lab of Philips discussed further in this session.

Field studies also present a credible, though challenging, approach to evaluating user experiences. Techniques such as diaries and experience sampling are recently extended to include technology support for facilitating self-report or enhancing them with the use of instrumentation like sensors, microphones and cameras.

This symposium brings together some leading researchers in this field, who will discuss their own efforts to meet the challenges discussed above. In all presentations, the emphasis is on methodology rather on the results of any specific study.

Symposium contents

Measuring fun in the home
Andrew Monk and Siân Lindley

Measuring urban mobility and encounter
Eamonn O'Neill and Vassilis Kostakos

Application driven experience research
Boris de Ruyter

InHaus-2: An innovative testbed for developing and testing ambient assisted living solutions and systems
Edwin Naroska

Probing in the wild: Lessons learned for contextual research
Manfred Tscheligi

Reading the tea-leaves in an intelligent Coffee Corner: understanding behavior by using sensory data
I. Mulder B. Hulsebosch, G. Lenzini, and M.S. Bargh

Measuring fun in the home

[1]Andrew Monk and [2]Siân Lindley
[1]Department of Psychology, University of York, UK, A.Monk@psych.york.ac.uk
[2] Microsoft Research, 7 J. J. Thomson Avenue, Cambridge, UK, v-silind@microsoft.com

Conventional measures of usability generally assume a task of some kind. One can then measure time to completion, errors, learning and so on. However, much of what we do in the home in purely for enjoyment. What is the task of watching the TV and what would an error look like? Is a game that can be learned faster than another better? Nevertheless, we may still want to construct prototypes and compare them in experiments for their ability to support enjoyment.

In the studies that motivated this talk [1, 2], we wished to compare the social affordances of different displays that might be used when people get together to share photographs. Are people likely to have a more enjoyable experience if control is equally distributed by giving everyone a remote control, for example? In order to make these experiments as natural as possible, they were conducted using groups of three friends who knew each other well. They were viewing their own photos, including photos of events they had all experienced and they were sharing the photos in the York Responsive Home, a laboratory set up to provide a home-like environment in a three-bedroom bungalow on the university campus.

The most common way of measuring enjoyment is with a post-experience questionnaire. However, Lindley and Monk [2] suggest that post-experience questionnaires are a form of 'recounting' [3] rather than an indication of unfolding experience. That is, responses to these questionnaires, while being based on the memory of the experience, are conditioned by who is asking the question and why. What is needed is a direct measure of enjoyment that does not interrupt the experience and does not depend on memory. Work by Monk and Reed [4] identified a phenomenon they term conversational flow. Their Conversational Analysis of transcripts of recreational telephone conferences involving five or more participants yielded the following definition.

"Flow is spontaneous, relaxed, inclusive and flowing. It is composed of small turn utterances in close succession (latched talk), sometimes overlapping, but when this happens there is no competition. Everybody gets to take part. Topics are light and topic change occurs easily. There are no awkward silences."

Similar characterisations of enjoyable conversation have been noted by Edelesky [5] and Goffman [6]. What it suggests is that group behaviour may be taken as an indicator of positive user experience. This measurement model is expressed as a hypothetical causative path diagram in Figure 1. In this model, social affordances encourage certain sorts of individual behaviour as well as certain emergent group behaviours. For example, Lindley and Monk [2] suggest a measure of the equality of a conversation as an indicator of enjoyment. This is because there is a tight feedback loop between such emergent measures of group behaviour, the unfolding experience of individual participants and their behaviour. Post experience questionnaires on the other hand measure something else, termed here recounted experience.

The talk will describe some of the measures we have used, including: equality, freedom, number of turns and turn overlap. These are all derived from analyses of speech in video records using The Observer®. In addition it will make a case for the validity and reliability of these measures.

References

1. Lindley, S. E., Monk, A. F. (2006) Designing appropriate affordances for electronic photo sharing media, in: *Extended abstracts of the SIGCHI conference on human factors in computing systems*. ACM Press, New York, 1031-1036.
2. Lindley, S. E. and Monk, A. F. (submitted) Social enjoyment with electronic photo displays: Awareness and control. Accepted subject to revision, *Int. J. Human-Computer Studies*.
3. McCarthy, J., Wright, P. (2004) *Technology as experience.* MIT Press, Cambridge MA.
4. Monk, A.F., Reed, D.J. (2007) Telephone conferences for fun: Experimentation in people's homes, in: *Proceedings of Home Informatics and Telematics*. Springer, New York, 201-204.
5. Edelsky, C. (1981) Who's got the floor? *Language in Society* **10**, 381-421.
6. Goffman, E. (1961) Fun in games, in *Encounters: two studies in the sociology of interaction*, Goffman, E. (ed) Bob Merril: Indianapolis.

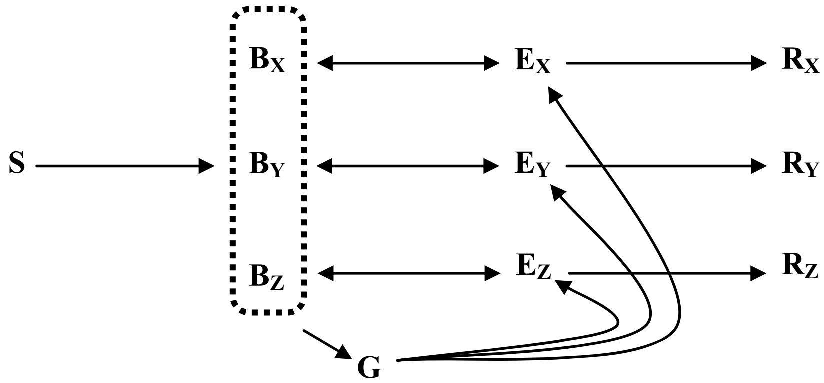


***Figure 1**. Causative model showing links between environment, behaviour and experience. S = social affordances of technology and physical context; B_i = behaviour of individual i; G = group behaviour (e.g., conversational equality); E_i = unfolding experience of individual i; R_i = recounted experience of individual i (e.g., rating of felt fun).*

Measuring urban mobility and encounter

[1]Eamonn O'Neill and [2]Vassilis Kostakos
[1]*Department of Computer Science, Department of Computer Science, University of Bath, Bath BA2 7AY, UK, eamonn@cs.bath,.ac.uk*
[2]*Vassilis Kostakos, Dept of Mathematics & Engineering, University of Madeira, Madeira, Portugal, vassilis@cmu.edu*

In [O'Neill et al., 2006] we described our development of novel methods for systematically observing and recording patterns of pedestrian mobility and encounter in the city. As a central part of our approach, we automated the capture of longitudinal data on mobility and encounter of mobile Bluetooth devices. Our recording of large-scale longitudinal data allows us to make two significant advances beyond traditional approaches to measuring behaviour in the city. First, we can inform aggregate level modelling and analyses with real world empirical data. This should help us to validate and improve upon the often simple approaches to such modelling. Secondly, we can investigate and analyse data that relate to a single user or a specific group of users, thus individualising our analysis in ways not possible with traditional aggregate approaches.

With no central servers to facilitate communication, Bluetooth devices rely on a discovery protocol to identify nearby devices. This protocol requires the initiating device to carry out an inquiry scan in a specific range of frequencies and wait for nearby devices to advertise their presence by transmitting their unique identifier. Thus, each inquiry scan provides information about which devices are in range at a discrete point in time. In our data collection we make use of the three key characteristics of Bluetooth: physical proximity, the explicit advertisement of the device's presence, and the unique identifier transmitted by each device.

Our combined methodology of manual and Bluetooth gatecounts allows us to estimate the penetration of discoverable Bluetooth in the urban population. As reported in [O'Neill et al., 2006], we found that for the city of Bath approximately 7.5% of observed pedestrians had discoverable Bluetooth devices. Our more recent measurements show a dramatically higher absolute number of Bluetooth devices. We are currently planning a new round of our combined observation methodology to investigate if this reflects a higher proportion of Bluetooth activity amongst the urban population.

The data record of Bluetooth activity is fundamentally a set of individual Bluetooth discovery events. In making sense of these data, we need to relate the individual events to a particular device and to its patterns of presence and absence across given scanner sites. In investigating encounter, we also need to relate these patterns across different devices. A temporal view allows us to begin making sense of the individual Bluetooth discovery records. Because of the use of unique identifiers in the Bluetooth protocol, each device can be associated with one and only one timeline across all our scanning locations in the city. A device moving past a scanner will generate a series of successive contact points on its timeline. Visualisations of the timelines reveal patterns of transience and persistence varying across times and spaces in the city and allow us to begin relating characteristics of those differing times and spaces to these data patterns. These data visualisations provide the foundation for an approach to making sense of our data in terms of three distinct abstractions: sessions, encounters and trails.

A session is defined as a set of contact points having no more than a threshold temporal distance $\delta 1$ between any two consecutive points. Thus, a session has an associated device, a start time, duration, and an associated location in the city (i.e. the scanner site). In the work reported in [O'Neill et al., 2006] we empirically derived appropriate values for $\delta 1$ by correlating human observations with Bluetooth observations. The concept of a session is central to our analyses, since it gives a time dimension to the discrete contact points generated by our scanners. Our next concept, encounter, builds on the concept of session. Encounter describes instances when two devices have been copresent. Thus, an encounter is defined by two devices, a location, a starting time and duration. To detect encounters we look for temporally overlapping sessions that took place at the same location. Our final concept, trail, extends the concept of a session with the spatial dimension. A trail is defined as a set of consecutive sessions for a given device, having no more than a threshold temporal distance $\delta 2$ between any two consecutive sessions. A trail, therefore, has an associated device, starting time, duration and number of hops (number of distinct sessions). Once again, $\delta 2$ has been empirically derived, and is based on our knowledge of the typical journey times between the physical locations we are observing.

In making sense of the patterns of movement and interaction of devices and people around the city, we first consider the distribution of session duration across our different scanning sites. We distinguish between persistent and transient devices using a threshold for session duration of 90 seconds. We empirically derived this threshold by measuring the session duration for individuals who walked past our scanners at a comfortable walking speed. This threshold of around 90 seconds allows us to establish empirically a conceptual distinction between transient and persistent devices and we can study how each conceptual group appears in different urban spaces.

We can represent trails as directed paths across a network graph. Each node can have metadata associated with it, such as duration of session, related semantic information (e.g. name, location co-ordinates, and so on), the identifiers of the devices that have visited it and various computed statistics such as frequency and average session duration. Thus, by preserving all the information recorded by each individual trail we can begin to analyse and compare trails. Graphs offer an effective way to inspect a set of trails and explore the relationships amongst them. For any given set of trails matching a set of criteria, we are able to inspect their layout and identify patterns. For instance, searching for the most popular trails late on a Friday night we can identify the taxi ranks as being the destination for many trails.

Visualising and analysing our raw Bluetooth activity data as sessions and trails allows us to begin making sense of the data in terms of people's behaviours in various forms of urban space (such as contrasting patterns of persistence between the pub and the street). Associating a unique timeline with every newly discovered device also allows us to trace the progress, or trail, of a device (and its user) by analysing the device's sequential presence at different scanning sites. A third crucial aspect of investigating the relationships between people, technologies and the city directly links the temporal and the spatial. Copresence or encounter requires that 2 or more devices are in the same space at the same time. It is in encounters that interactions occur: interactions between person and person, between person and fixed device, between mobile device and mobile device, between mobile device and fixed device, and so on. To study the patterns of encounter in the city, we first identify device sessions that overlap in time and were recorded at the same location.

Again we can represent these patterns as a network graph. Assuming that each device from our dataset becomes a node in this graph, the list of encounters describes the links between all nodes. Thus, we are able to generate social network graphs [e.g. Strogatz, 2001] that represent the patterns of encounter across our entire dataset. We can generate various graphs from our data, such as an individual social network graph per scanner site, or a graph of our entire dataset in one city-scale social network graph. Furthermore, we can generate these graphs over the entire lifetime of our scanning or over any specified period. An array of standard metrics such as closeness, clustering coefficient, etc [Freeman, 2004] enables us to identify meaningful subgroups of the population, and focus our attention on them. For instance, we might identify isolated individuals and examine the trails these individuals take across the city.

The approach presented here allows us to begin making sense of the data derived from measuring urban behaviour temporally and spatially. It reveals patterns of transience and persistence varying across times and spaces in the city and allow us to begin relating characteristics of those differing times and spaces to the data patterns. In related work, we have moved from visualisations to the development of more formal and systematic analytical concepts and tools that allow us to automate aspects of our data analysis.

References

1. Freeman, L. (2004) The Development of Social Network Analysis: A Study in the Sociology of Science. Vancouver, BC, Canada: Empirical Press.
2. O'Neill, E., Kostakos, V., Kindberg, T., Gen. Schieck, A.F., Penn, A., Fraser, D.S., And Jones, T. (2006) Instrumenting the City: Developing Methods for Observing and Understanding the Digital Cityscape. In *Proceedings of UbiComp* 2006, 315-332.
3. Strogatz, S.H. (2001) Exploring complex networks. *Nature*, **410**, 268-276.

Reading the tea-leaves in an intelligent Coffee Corner: understanding behavior by using sensory data

I. Mulder[1,2], B. Hulsebosch[1], G. Lenzini[1], andM.S. Bargh[1]
[1] *Telematica Instituut, Enschede, the Netherlands, ingrid.mulder@telin.nl*
[2] *Institute for Communication, Media, and Information Technology, Rotterdam University, the Netherlands*

Abstract

This article discusses the challenges for methodological innovation based on experiences in an experimental "Living Lab" setting; an interactive and context-aware coffee corner in a research institute where people that take coffee can use a variety of services offered by intelligent environment at the coffee corner's site. It also collects sensory information of users while they are moving along the institute and when they interact with the coffee corner. The analysis of contextual data allows the construction of a behavioral model of users in a non-intrusive manner. We explain how this method can be used to get insight in measuring behavior in context in an unobtrusive way.

Keywords

Living Lab, Context-Awareness, Behavioral Models, sensory data

Introduction

Human-centered design is nowadays a common practice and many methods and tools are at hand to this scope. However, we still see many projects that face difficulties in designing intelligent systems that meet (future) users' needs. The reason might be emerging technology itself, because we face more complexity in designing intelligent environments. Nevertheless, the exploitation of such environments enables researchers to come close to the users and to understand their expectation when using those environments.

A Living Lab exploits intelligent infrastructures for measurement purposes instruments, moves research out of laboratories into real-life contexts, and provides opportunities to non-intrusively study social phenomena in users' social and dynamic context of daily life. The Living Lab concept has been acknowledged in Europe as an open innovation instrument that is appropriate to study questions related to human behavior and experiences; involving areas of user interface design and ergonomics as well as user acceptance, extending to user co-design process, and it leads to service or product creation [9]. It might be clear that the Living Lab concept opens a wealth of possibilities to exploit the evaluation of intelligent environment. However, according to Mulder and Kort [7] "many of the automated tools alone do not deliver the desired insight; they need to be combined with common methods such as interviews and focus groups which either provide input for the automated measurements (which things should be captured and asked for during experience sampling) or provide additional information after the automated measurements (clarifications of specific experience sampling data, behaviors or contexts in which it appeared)".

Differently stated, there is still a need for research in methodological guidelines and tool requirements for data-analysis. In particular, analysis techniques are required for correlating objective behavior and subjective user experience data into relevant design context parameters.

In the remainder of this article, we describe an experimental Living Lab setting to get insight in measuring behavior in context and in an unobtrusive way. Using a context-management framework infrastructure, we are able to collect, store, and analyze a great amount of contextual data. The analysis of contextual data allows the construction of a behavioral model of users in a non-intrusive manner. Starting from this experimental setting, issues for data collection and analyses are discussed, as well as the current availability of methods and tools for building and exploiting user behavioral and experience models are reviewed in general Living Lab' scenarios.

The Intelligent Coffee Corner

The intelligent Coffee Corner is a real-life coffee space with reasoning capabilities and intelligent services located throughout a research institute, which employs about a hundred workers situated in two connected buildings. Each building has four floors. Moreover, the employees that work in different projects are spread (rather randomly) across different office locations. Every floor has a coffee space and is equipped with a high density of sensors allowing for device discovery and human detection by using Bluetooth dongles, RFID readers, WLAN access points, video cameras, pressure mats, computers, and advanced displays (Figure 1).

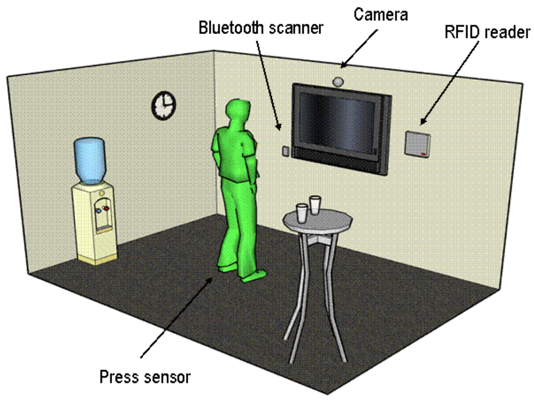


Figure 1. *The intelligent Coffee Corner, equipped with sensors such as Bluetooth dongles, RFID readers, WLAN access points, video cameras, pressure mats, computers, and advanced displays.*

Most employees carry detectable devices (e.g., Bluetooth-enabled mobile phones or PDAs and WLAN-enabled laptops) with them. In addition, all employees wear by default a RFID-enabled badge, which is needed to open doors in order to access the different floors in the building. These badges are also used to sense employee locations throughout the institute. Similar to the Living Lab concept, our intelligent Coffee Corner finds success if people and technology continually interact [8].

Coffee break brainstorming, questionnaires, surveys, scenarios, contextual inquiries, participatory design, focus groups, paper prototypes, in depth interviews, and technology probes are some of the traditional tools that have been used so far to gather user needs, expectations, perspectives, ideas, feedback, or inspiration during users' daily activities [2]. While most of these tools engage users' involvement in an explicit way, our intelligent Coffee Corner also provides tools to get information from users and insight in their behavior in an implicit and less obtrusive manner. Examples of this implicit way of data collection that uses the capabilities of intelligent environments are listed in the following:Logging: information about the use of application, which and how frequently.

Sensing: information about the (physical or virtual) context of the user.

User-generated content: information that users have created on their own initiative; thus for other reasons than research might have intended.

Sensing of contextual information

For obtaining information about users' behavior and users' experience we make use of an infrastructure that enables the collection and management of heterogeneous context information obtained from various heterogeneous sensors. Such an infrastructure is called Context Management Framework (CMF) whose design is described in [10].

The CMF is a highly distributed service infrastructure that enables context-sensitive applications to discover and obtain context information. Examples of context information supplied by the CMF include GPS location coordinates, WLAN access point associations, RFID reader data, Bluetooth scan data, desktop keyboard typing status, presence information, and Outlook calendar meetings. In an implicit way, the CMF collects raw data from these context sources, processes the data by fusion and reasoning to infer higher-level context information and/or better quality context information so that it is useful for a service provisioning that meets user requirements. Inference in the CMF is done with various reasoning components that fuse and enrich sensed information to higher semantic levels. Each reasoning component can use its own internal algorithm and inference mechanism.

To have a shared understanding of the meaning of the information that is delivered and exchanged by the CMF, the sources of information exchange their information as instances of a shared ontology. The (extensible) ontology describes both the types of context information as well as the relations between these types. An overlay framework takes care of aggregating context information per entity. It does so by using specialized broker components that, in addition, are able to enforce policy rules. This allows users for instance to specify who may access their privacy-sensitive context data [3].

In summary, thanks to this CMF we are able to collect, store, and analyze a great amount of contextual data and opens possibilities for advanced study of user experience and user behavior. The CMF has proven to be a robust and flexible underlying infrastructure for several mobile health and office applications. One such example is the "Colleague Radar" application, a location-tracking service with context-aware security and privacy features. The "Colleague Radar" enables access to context of colleagues, among which, their location inside or outside the office. A screen shot of the application user interface (as it runs in the Coffee Corner) is shown in Figure 2.

The Colleague Radar application is an example of presence-aware services whose benefit from an intelligent management of context-information is high. Its identification and authentication mechanism is designed to be dependent on contextual information. User authentication is done in a very user friendly manner based on the locations of different user identity tokens such as an RFID employee card or Bluetooth phone, that the user is assumed to carry with him/her [4]. This is supplemented with face recognition to improve user identification and authentication even more [5].

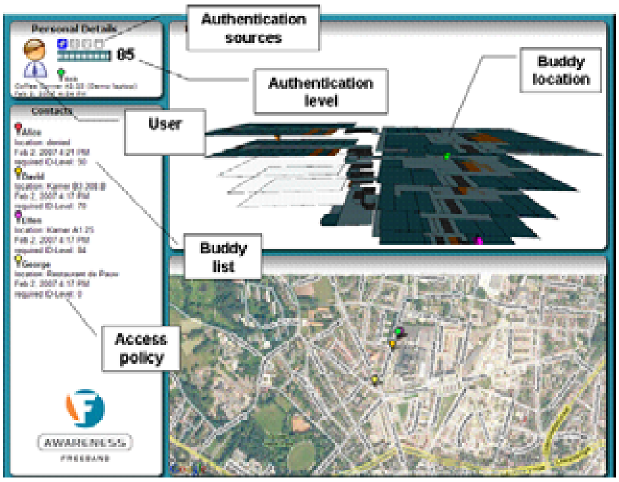


Figure 2. *A screen shot of the Colleague Radar application. The up-left panel shows the ID-level of the user and the devices that have been used in the identification process. The down-left panel is the list of colleagues that have agreed to have their position shown considering the user's identity and current ID-level. The two right panels show, graphically, the position of the colleagues in the building, and in the area around.*

The availability of advance type of contextual data, such as behavioral models tailored per user, would help in improved user identification and authentication procedures (users can be recognize thanks to their behavioral patterns) whose reliability is a key quality for the success of this kind of application.

Sensory data for measuring behavior

The high density of sensors and intelligent devices, managed through the CMF, enable not only to support device discovery and human detection and identification, but also to observe user behavior based on the sensed information. In this section we discuss the challenges for methodological innovation in understanding user behavior and user experience in context.

Context information, when aggregated in time, can potentially be used to determine behavioral models of users and, consequently, to improve the provisioning of customized services. One general methodology consists of processing raw data that originates from context sources and to obtain higher quality context information. For example, probabilistic methods can be applied to positional raw data to evaluate the expectation we have in a person's position. This measure can be useful to build for instance automatic authentication or authorization of users depending on the location of their identity tokens. When this information is combined with other kinds of contextual data, for example location or activity information obtained from Outlook agenda, even more accurate reasoning can be done; for example the use of a plotter can be allowed in a certain meeting room only when the user is actually in the room and attending to a meeting.

The collection and the organization of raw contextual data into an event structure (e.g., a temporal timeline) allow for the understanding of advanced situational events. By using model checking algorithms for temporal logics (cf. [1]) it is possible to check properties that express user behavior in specific situations, like for example, "might the user have forgotten the meeting?". Such an event, in fact, can be translated into temporal expressions like "the user has the meeting in his Outlook agenda, which he usually synchronizes with his mobile phone and he is actually in the library". The recognition of contextual and temporal situations fosters the design of better and innovative services. In the previous example, the missing user could be automatically alerted of the meeting on his mobile phone (despite the fact that his local agenda is not updated); if he has forgotten the device somewhere, a close colleague can be contacted instead.

The evaluation of temporal events can be either momentary (e.g., where he is now), or can be used reconstructive (e.g., where he has been) or anticipative (e.g., where he is likely to go). This last aspect is related to prediction of events, which is quite difficult but still possible. Statistical model, constructed from existing sensory data, can be utilized to predict a person's indoor location at a specific point in time. Data mining techniques based on machine learning methods (e.g., Bayesian Networks, neural networks, decision trees) and plan recognition techniques are available. Reasoning on (not strictly temporal) context information aims to improve the prediction. Bluetooth dongles and desktop activity as well as Outlook Calendar information can be used to determine and predict the user's movements.

The analysis of context information allows also an unobtrusive identification. When a user is trying to access a resource, his identity is estimated by processing the contextual information gathered so far, for example by processing the position of the devices that are expected to be carried by the authenticating identity. Automatic identification can be particularly useful when privacy is also requested to be preserved. First, contextual information can be preferred to other confidential identity token like passwords or PINs (the intelligent Coffee Corner, for example, is a public space potentially insecure to eavesdropping). Then the single contextual pieces of information remain related to anonymous (distinct) entities until an identification request is forwarded; in that moment, the system is allowed to calculate which identity is mostly believed trying to log in.

The use of contextual information allows also reasoning about the trustworthiness of certain actions. Inconsistencies in the event structure might be witnesses of untrustworthy situations. For example, a user mobile that is moving in a different direction from a user badge may mean that the user's phone has been stolen. In a city-wide scenario, a user withdrawing money from closer but different cash dispensers in a very short time (when the user is expected to be at work) may identify a debit card robbery. Temporal properties satisfiability could be enhanced with quantitative methods for trust evaluation; critical situation being labeled with a value indicating the level of criticality which is obtained by considering also past experiences or even recommendations.

Discusssion and conclusions

The analysis of contextual data for the inference of users behavior or movement patters is affected by a number of issues that any Living Lab must carefully take into account. Hereto, the following issues are critical:

Reliability and availability of context info. Are contextual data always reliable? What is some information is not available when needed?

Fusing different heterogeneous context sources: How to map patterns for each source to each other? How to fuse non-homogeneous data (e.g., a pattern in a location-database with one in a temperature database to derive someone is being ill or is running?)

Triggers for behavior. What does trigger a user to behave in a certain manner?

Effectiveness of sensory data for user behavior modeling: How much does the derived user behavior model match the real behavior? How does the fitness of a behavior model relate to the set of sensors used and their quality of context?

Data reliability is a first prerequisite for answering research questions. In observational research, reliability of data refers to the degree of agreement between sets of data collected independently from the same scene by two different sensors or by the same sensor at different times in the data collection process. Various quantitative measures have been used by researchers for the assessment of the degree of agreement between sensors or observers. Jansen, Wiertz, Meyer and Noldus [6] discuss several methodological problems related to the assessment of observer agreement on observational data, how these can be solved, and how these solutions have been implemented?

Besides human behavior, the characteristics of unobtrusively observing a person in a real-world environment should be taken into account. Some physical or virtual phenomena are difficult or impractical to observe due to the availability, cost or obtrusiveness of the required sensor. Observations may be missed due to sensor hardware failures, connectivity problems or the user moving outside the coverage range of a sensor. In addition, the quality of an observation depends on the characteristics of the sensor, such as its accuracy and sensitivity. Finally, observations of different phenomena are often related. In the reasoning process, the qualities of each observation as well as relationships between observations thus need to be taken into account.

Typical approaches for determining user behavior assume the availability of actions or derive actions from sensor data in a single step. In contrast, we argue that inferring actions from sensor data in multiple steps is more effective. This multi-step approach allows for specialized reasoning techniques to tackle specific parts of the inference process. At various abstraction levels, contradictions and superfluities can be eliminated and missing observations can be compensated by combining and interpreting contextual information. This process results in a gradual reduction of the observation space while enriching the context information. The multi-step processing of sensor data into actions and enriched context information is facilitated by the described CMF.

The availability of context information gathered from multiple sensor sources provided unprecedented opportunities to study user behavior and experience in a non-intrusive, natural manner. Finally, we are able to study the impact of innovative and intelligent solutions in a naturally way that is not intrusive for users. However, a lot of work needs to be done in the area of context reasoning and behavior assessment.

Acknowledgements

This paper has been inspired by discussions in the Intelligent Communication team and the work done earlier in the project "Freeband Awareness" which is part of the research programme Freeband Communication (www.freeband.nl). We thank our colleagues that have contributed to the Coffee Corner research and the Colleague Radar in particular.

References

1. Clarke, E.M.; Grumberg, O.; Peleg, D.A. (2000). *Model Checking*, MIT Press.
2. Fachry, K.; Mulder, I.; Eertink, H.; Janse, M. (2007). Enticing Sociability in an Intelligent Coffee Corner. In G. Qiu; C. Leung; X. Xue; R. Laurini (eds.). *Advances in Visual Information Systems*, *9th International Conference*, VISUAL

2007 Shanghai, China, June 28-29, 2007, Revised Selected Papers, LNCS 4781 (pp. 490-501). Springer: Berlin/ Heidelberg.

3. Hesselman, C.; Eertink, H.; Wibbels, M. (2007). Privacy-aware Context Discovery for Next Generation Mobile Services. In *Proc. of the 3rd Workshop on Next Generation Service Platforms for Future Mobile Systems (SPMS 2007)*, SAINT2007, January 15-19, 2007, Hiroshima, Japan.
4. Hulsebosch, R. J.; Bargh, M.S.; Lenzini, G.; Ebben, P.W.G.; Iacob, S.M. (2007). Context sensitive adaptive authentication. In *Proc. of the 2nd European Conference on Smart Sensing and Context*, Kendal, UK, 23-25 October, 2007, LNCS 4793 (pp. 93-109), Springer: Berlin/ Heidelberg.
5. Hulsebosch, R.J.; Ebben, P.W.G. (2008). Enhancing Face Recognition with Location Information. In *Proc. of "ARES 2008 – The International Dependability Conference"*, the Third International Conference on Availability, Reliability and Security, March 4th-7th, 2008, Barcelona, Spain.
6. Jansen, R.G.; Wiertz, L.F.; Meyer, E.S.; Noldus, L.P.J.J. (2003). Reliability analysis of observational data: Problems, solutions, and software implementation. *Behavior Research Methods, Instruments & Computers*, **35**, 391-399.
7. Mulder, I.; Kort, J. (2008). Mixed emotions, mixed methods: the role of emergent technologies to study user experience in context. In S. Hesse-Biber; P. Leavy (eds.). *Handbook of Emergent Methods in Social Research* (pp. 601-612). New York: Guilford Publications.
8. Mulder, I.; van Kranenburg, H.; Velthausz, D. (2007). The Intelligent Coffee Corner: an innovation experimentation space. In P. Cunningham; M. Cunningham (eds.). *Expanding the Knowledge Economy: Issues, Applications, Case Studies* (pp. 781-788). IOP Press.
9. Mulder, I.; Velthausz, D.; Kriens, M. (2008). Living Methodologies: Understanding the Dynamics of Innovation. In J. Schumacher; V-P. Niitano (eds.). *European Living Labs – a new approach for human centric regional innovation* (pp. 31-38). Berlin: Wissenschaftlicher Verlag Berlin.
10. van Kranenburg, H.; Bargh, M.S.; Iacob, S.; Peddemors, A. (2006). A Context Management Framework for Supporting Context Aware Distributed Applications. *IEEE Communications Magazine*, **44(8)**, pp. 67-74.

Clinical and preclinical models for depression Symposium

H.W.M. Steinbusch
European Graduate School of Neuroscience and University of Maastricht, Maastricht, The Netherlands, h.steinbusch@np.unimaas.nl

Depression–like behaviour is both in humans as well as in rodents difficult to measure. In this symposium the speakers will cover the areas of measuring the differences between humans, rats and mice in anxiety-related behaviour and its relation towards cognition and stress. Finally, two speakers will deal with the effects of the immune-system on anxiety-like behaviour and its possible implications towards further neurodegeneration.

Symposium contents

Prenatal stress produces anxiety- and depression-related behavior particularly in male Sprague-Dawley rats

D.L.A. Van den Hove, G. Kenis, M. Bruschettini, C.E. Blanco, H.W.M. Steinbusch and J. Prickaerts

Synaptic plasticity dysfunction in vivo

T. Ondrejcák, B. Ryan, I. Klyubin, W.C. Cullen, and M.J. Rowan

Interferon alpha induced depression-like behavior in the rat

Aye Mu Myint, Harry W.M. Steinbusch, and Brian E. Leonard

The olfactory bulbectomised rat as a model of depression

Brian E. Leonard

Measuring behavior with chronic stress depression models in mice

T. Strekalova

Prenatal stress produces anxiety- and depression-related behavior particularly in male Sprague-Dawley rats

D.L.A. Van den Hove[1*], G. Kenis[1*], M. Bruschettini[1], C.E. Blanco[2], H.W.M. Steinbusch[1], J. Prickaerts[1]

[1]Department of Psychiatry and Neuropsychology, Research Institute Brain and Behavior, Division Cellular Neuroscience, Faculty of Medicine, European Graduate School of Neuroscience (EURON), Maastricht, The Netherlands {d.vandenhove@np.unimaas.nl}

[2]Department of Pediatrics, Research Institute Growth and Development, Faculty of Medicine, Maastricht University, Maastricht, The Netherlands

**These authors contributed equally to this work*

Background

Nowadays, it has become increasingly clear that adverse events experienced by a pregnant woman, be they physical or emotional, may be reflected upon the developing fetus and adversely affect its physical and mental wellbeing in later life. In the present study we examined the effects of prenatal restraint stress on anxiety- and depression-related behavior in both male and female Sprague-Dawley rats at 4-5 months of age.

Behavioral assessment

Anxiety-related behavior was evaluated in the elevated zero maze and the home cage emergence test. Depression-related behavior was examined using the forced swim test

The elevated zero maze consisted of a circular alley (diameter of 100 cm; path width 10 cm) made from black plastic material that was transparent for infrared light and elevated 20 cm above the floor. The maze was divided in four parts, i.e., two opposite open parts and two opposite closed parts with sidewalls 30 cm height. The open parts had borders with a height of 5 mm to prevent the rat from stepping down from the apparatus. For the test, the rat was placed into one of the open parts facing a closed part of the apparatus. After five minutes the rat was removed from the apparatus and the maze was cleaned with ethanol and water and dried thoroughly. The movements of the rat were scored automatically under dark conditions with a computerized system using an infrared video camera (EthoVision Pro, Noldus, The Netherlands). Percentage of time spent in the open part of the maze and total distance traveled were determined.

For the home cage emergence test the rat's home cage (opened) was placed in the center of an open field ($1m^2$) and the rat was allowed to leave its cage via a grid walkway. The latency to emerge from the home cage (i.e., four paws on the grid) was scored. If the rat did not emerge from its home cage within 300 sec, the session was ended, the home cage was closed again and the rat was given a score of 300 sec. This test was carried out on 3 consecutive days under low light conditions.

In the forced swimming test four cylindrical glass tanks (50 cm tall, 20 cm in diameter) were filled to a height of 30 cm with 25°C water. The movements of the rat were scored automatically with a computerized system (EthoVision Pro, Noldus, The Netherlands) during a 5 min session under low light conditions. Scored were 'immobility' which reflects no movement at all and/or minor movements necessary to keep the nose above the water) and 'strong mobility' reflecting 'escape behavior' (e.g. climbing against the walls and diving). Settings within EthoVision were adjusted based on manually recorded sessions and were attuned for each gender separately (immobility/mobility threshold: 12 and 20; mobility/strong mobility threshold: 16.5 and 23.9 for males and females, respectively).

Results

Prenatal stress was associated with a clear increase in anxiety-related behavior in male, but not female offspring, as evaluated in the elevated zero maze and the home cage emergence test. Likewise, depression-related behavior in the forced swim test was increased in prenatally stressed male rats only. Prenatally stressed male offspring further showed increased basal plasma corticosterone levels, whereas both prenatally stressed males and females failed to show an adequate response to stress with lower stress-induced corticosterone levels as compared to controls. Female hippocampal weight was relatively higher after prenatal stress, which may explain the absence of clear behavioral effects of prenatal stress in this gender. In addition, male birth weight was a predictive marker both for performance in the forced swim test, as well as for plasma corticosterone levels in adulthood.

Discussion

In conclusion, prenatal stress resulted in increased anxiety- and depression-related behavior particularly in male Sprague-Dawley rats. Females seemed to be relatively protected towards the effects of prenatal stress.

Synaptic plasticity dysfunction *in vivo*

T. Ondrejcák, B. Ryan, I. Klyubin, W.C. Cullen, and M.J. Rowan
Department of Pharmacology and Therapeutics, Trinity College, Dublin 2, Ireland, mrowan@tcd.ie

Persistent activity-dependent increases and decreases in synaptic transmission provide an attractive biologically plausible means of information storage in the brain. Since different types of stress can facilitate or impair memory great interest has focused on how stress affects synaptic plasticity. Dysregulation of the induction or maintenance of hippocampal synaptic plasticity may be of importance in stress-related psychiatric disorders such as depression where normal hippocampal function is often impaired [1].

Exposure to a novel inescapable environment that causes behavioural freezing blocks the induction of long-term potentiation (LTP) of excitatory transmission in the rat hippocampus (CA1 area) [2-5]. It also can facilitate long-term depression (LTD) [2, 3]. The stress-induced switch in the direction of plasticity was found to be dependent on corticosteroid action at glucocorticoid receptors, and is regulated by drugs that alter brain 5-HT at least partly via 5-HT2 receptors. In contrast, exposure to a novel non-stressful inescapable environment that causes active exploration triggers a rapid reversal of previously established LTP within a defined time window but also facilitates new LTP induction in a different time window [6, 7].

Somewhat similar to the effects of inescapable stress, we have discovered that hippocampal LTP is inhibited by the Alzheimer disease-related peptide amyloid ß (Aß) in vivo after intracerebroventricular injection [8-12]. Moreover Aß also promotes LTD in vivo [13]. Such disruption of plasticity mechanisms is mediated by small abnormally folded oligomers of Aß via pro-inflammatory/ oxidative and nitrosative stress mechanisms and is dependent on many neurotransmitter receptors / signaling pathways and integrins [14].

Here we present recent data comparing the effects of treatments that disrupt synaptic plasticity in vivo either in awake, freely behaving animals or under anaesthesia. Briefly, Male Wistar rats are housed under a twelve h light-dark cycle. Prior to surgery animals were anaesthetised with urethane (ethyl carbamate; 2.1g/kg, i.p.) for non-recovery experiments or a variety of anaesthetics for recovery experiments. Monopolar recording electrodes and bipolar stimulating electrodes are made in the laboratory from two lengths of Teflon coated tungsten wire. A stereotaxic apparatus is used to place the electrodes in the CA1 area of the dorsal hippocampus. Electrophysiological criteria are used to determine the optimal electrode placement. Field excitatory post-synaptic potentials (fEPSPs) are recorded from the stratum radiatum following stimulation of the Schaffer collateral-commisural pathway. Test fEPSPs are evoked by a single square wave pulse of current at low frequency (0.033Hz). The test stimuli evoked responses of between 50-55% maximum fEPSP amplitude. The high frequency stimulation (HFS) protocol to induce LTP comprises trains of pulses at different frequencies (100-400Hz). For recovery experiments animals are habituated to the recording box and at least one week is allowed for recovery from surgery.

We have discovered many different effects of the different treatments on baseline synaptic efficacy and synaptic plasticity, depending on experimental protocol and behavioural conditions. The presentation will outline some of these experimental variables.

The ability of stress to dramatically alter the induction of different types of synaptic plasticity in the hippocampus is believed to play a key role in mediating the effects of stress on hippocampus-dependent memory and learning. Increasing our understanding of the contributions of behavioural factors and cellular mechanisms should greatly aid the development of new approaches to stress-related cognitive disorders.

Supported by Science Foundation Ireland, Irish Health Research Board, the Irish Higher Education Authority (PRTLI) and the EU.

References

1. Campbell, S. and Macqueen, G. (2004) The role of the hippocampus in the pathophysiology of major depression. *J Psychiatry Neurosci*. **29**, 417-426
2. Xu, L., Anwyl, R. and Rowan, M. J. (1997) Behavioural stress facilitates the induction of long-term depression in the hippocampus. *Nature*. **387**, 497-500
3. Xu, L., Holscher, C., Anwyl, R. and Rowan, M. J. (1998) Glucocorticoid receptor and protein/RNA synthesis-dependent mechanisms underlie the control of synaptic plasticity by stress. *Proc Natl Acad Sci U S A*. **95**, 3204-3208
4. Shakesby, A. C., Anwyl, R. and Rowan, M. J. (2002) Overcoming the effects of stress on synaptic plasticity in the intact hippocampus: rapid actions of serotonergic and antidepressant agents. *J Neurosci.* **22**, 3638-3644
5. Ryan, B. K., Anwyl, R. and Rowan, M. J. (2008) 5-HT(2) receptor-mediated reversal of the inhibition of hippocampal long-term potentiation by acute inescapable stress. *Neuropharmacology* in press.
6. Li, S., Cullen, W. K., Anwyl, R. and Rowan, M. J. (2003) Dopamine-dependent facilitation of LTP induction in hippocampal CA1 by exposure to spatial novelty. *Nat Neurosci*. **6**, 526-531
7. Xu, L., Anwyl, R. and Rowan, M. J. (1998) Spatial exploration induces a persistent reversal of long-term potentiation in rat hippocampus. *Nature.* **394**, 891-894
8. Cullen, W. K., Suh, Y. H., Anwyl, R. and Rowan, M. J. (1997) Block of LTP in rat hippocampus in vivo by beta-amyloid precursor protein fragments. *Neuroreport*. **8**, 3213-3217
9. Klyubin, I., Betts, V., Welzel, A. T., Blennow, K., Zetterberg, H., Wallin, A., Lemere, C. A., Cullen, W. K., Peng, Y., Wisniewski, T., Selkoe, D. J., Anwyl, R., Walsh, D. M. and Rowan, M. J. (2008) Amyloid beta protein dimer-containing human CSF disrupts synaptic plasticity: prevention by systemic passive immunization. *J Neurosci*. **28**, 4231-4237
10. Klyubin, I., Walsh, D. M., Lemere, C. A., Cullen, W. K., Shankar, G. M., Betts, V., Spooner, E. T., Jiang, L., Anwyl, R., Selkoe, D. J. and Rowan, M. J. (2005) Amyloid beta protein immunotherapy neutralizes Abeta oligomers that disrupt synaptic plasticity in vivo. *Nat Med*. **11**, 556-561
11. Walsh, D. M., Klyubin, I., Fadeeva, J. V., Cullen, W. K., Anwyl, R., Wolfe, M. S., Rowan, M. J. and Selkoe, D. J. (2002) Naturally secreted oligomers of amyloid beta protein potently inhibit hippocampal long-term potentiation in vivo. *Nature*. **416**, 535-539
12. Wang, Q., Klyubin, I., Wright, S., Griswold-Prenner, I., Rowan, M. J. and Anwyl, R. (2007) alphav integrins mediate beta-amyloid induced inhibition of long-term potentiation. *Neurobiol Aging* in press.
13. Kim, J. H., Anwyl, R., Suh, Y. H., Djamgoz, M. B. and Rowan, M. J. (2001) Use-dependent effects of amyloidogenic fragments of (beta)-amyloid precursor protein on synaptic plasticity in rat hippocampus in vivo. *J Neurosci*. **21**, 1327-1333

14. Rowan, M. J., Klyubin, I., Wang, Q., Hu, N. W. and Anwyl, R. (2007) Synaptic memory mechanisms: Alzheimer's disease amyloid beta-peptide-induced dysfunction. *Biochem Soc Trans.* **35**, 1219-1223.

Interferon alpha induced depression-like behaviour in the rat

Aye Mu Myint[1], Harry W.M. Steinbusch[2], and Brian E. Leonard[3]

[1] *Laboratory for Psychoneuroimmunology, Psychiatric Hospital, Ludwig-Maximilians University, Munich, Germany, ayemu.myint@med.uni-muenchen.de*

[2] *Division of Cellular Neuroscience, Department of Psychiatry and Psychobiology, Research Institute Brain and Behaviour, University of Maastricht, The Netherlands*

[3] *Department of Pharmacology, National University of Ireland, Galway, Ireland*

Background

One of the accepted pathophysiological mechanisms of major depressive disorder is activated inflammatory response system (IRS) which is indicated by increased production of pro-inflammatory cytokines. The interferon-α (IFN-α) is a pro-inflammatory cytokine used in the treatment of chronic hepatitis C infection and certain type of cancer that may and this treatment can induce depressive symptoms as one of the common side effects. Here, we discuss about the behavioural changes in IFN-α treated rats as a model of activated IRS.

Animals and behavioural assessment

Rats – Male Wistars (250-300gm) were assigned as 4 groups of 10: saline control - subcutaneous injection of 0.9% Saline (0.2ml/kg) daily for 7 weeks; paroxetine Control - paroxetine (10mg/kg,oral) for 7 weeks; Interferon (IFN) - saline for 14 days followed by subcutaneous injection of Interon A (IFN-α) (provided by Schering Plough) (50,000 IU/kg) 3 days/week for 5 weeks; paroxetine + Interferon (P+IFN) - paroxetine for 14 days followed by Interon A for 5 weeks.

Behavioural and Biochemical procedures – The behavioural tests were started two week after interferon. Open field test was performed in a square arena with 5 min session per day for four days. Morris water maze test was performed in a water tank (21°C) with invisible escape platform on 3 subsequent days with 8 trials a day and a probe trial after the last trial. Home Cage Emergence test was applied once a week. The latency periods to emerge from the home cage were recorded. Data recording and analyses for open field test and water maze were done using EthoVision software.

The pro-inflammatory IL1-β, TNF-α and anti-inflammatory IL-10 cytokines were analyzed in the prefrontal cortex, hypothalamus and hippocampus areas of the brain and supernatant from both unstimulated and stimulated whole blood culture.

Results

In open filed, the IFNα treated rats showed anxiety behaviour compared to the rats from the other groups. There was no significant difference in home cage emergence test, Morris water maze and object recognition test. There is no significant difference in plasma corticosterone between groups. The pro-inflammatory cytokines (TNFα, IL1β and IFNγ), were significantly higher whereas the anti-inflammatory cytokine, IL10 was lower in the stimulated whole blood culture of IFNα treated rats. In the brain, both pro-inflammatory cytokine IL1β and anti-inflammatory cytokine IL10 were higher in hypothalamus of the IFNα treated rats; by contrast the concentration of IL10 was lowest in hippocampus region of this group compared to the other groups. The paroxetine pre-treated rats did not show these behaviour and cytokine changes following IFNα treatment.

Conclusions

The pro- and anti-inflammatory cytokines could activate the immune response and induce behavioural changes and those changes were attenuated by the pre-treatment with SSRI, paroxetine. Open-field test might be the most sensitive behavioural test in measuring immune activation-related behavioural changes.

Measuring behavior with chronic stress depression models in mice

T. Strekalova
School for Mental Health and Neuroscience, Division of Cellular Neuroscience, Maastricht University, Maastricht, The Netherlands, t.strekalova@np.unimaas.nl

Chronic stress is regarded as one of the most adequate methods for mimicking depression in rodents, because it 1) fits the face validity criteria by implication of induction of anhedonia, a core symptom of depression, and 2) simulates etiological relevance of stress with this disorder [1, 2]. At the same time, studies, using chronic stress depression models, failed to define a consistent behavioral phenotype of a depressive-like state seen in chronically stressed rodents and overall, resulted in abstruse and variable outcome [3]. This greatly limits a value of this method in modeling depression and decreases its utility for the identification of new mechanisms / targets of antidepressant therapies. Studies with our stress-induced anhedonia model suggest that unresolved methodological difficulties in measuring behavior in chronically stressed animals may underlie these problems.

Lack of a proper control

Perhaps the most obvious problem with measuring behavior during stress-induced anhedonia consists in the shortcoming that all effects observed in groups of stressed animals are attributed to a depressive-like syndrome and anhedonic state. Meanwhile, stress *per se* can evoke a number of physiological alterations, which are not associated with depression. Since available chronic stress models did not provide a control for the effects of chronic stress alone, strictly speaking, it was not possible to relay findings obtained in chronically stressed animals selectively to a depressive-like phenotype and hedonic deficit. We established an anhedonia-evoking stress regimen, upon which not all individuals, but only about 50-70% of C57BL mice, develop a hedonic deficit [4]. Employment of the stressed non-anhedonic group as an internal control for the effects of chronic stress *per se* enabled the first attempt to separate behavioral features of anhedonia and consequences of chronic stress alone. We found that anhedonia is selectively associated with depressive-like behaviors, such as “behavioral despair” in the forced swim and tail suspension tests and decreased novelty exploration; these deficits were not observed in non-anhedonic animals. Increased anxiety and changes in locomotion were detected in stressed mice with and without hedonic deficit. Studies with a model of social defeat stress on the C57BL6 strain confirm our data on percentage of individuals resilient /susceptible to a development of stress-induced depressive-like state and on their behavioral differences [5].

Low resolution of sucrose test

Decreased intake of palatable solutions, e.g., of sucrose, is taken as a behavioral measure of hedonic deficit / depressive-like state [1]. Insufficient accuracy of the sucrose test in mice is another key difficulty in measuring behavior with chronic stress depression models. Typically, sucrose test can let to reveal the differences between the groups, but not between the individual mice. We showed that side preference, neophobia, individual differences in circadian patterns of liquid intake, experience of sucrose taste and other factors, when not taken under control, result in physiological and physical artifacts in evaluating sucrose drinking behavior [6]. Modification of this test that included switching the choice bottles during testing, allowed us to increase its resolution and assess the hedonic state of mice on an individual basis.

Behavioral artifacts resulted from a stress-induced hyperlocomotion

The majority of studies with chronic stress depression models demonstrated paradoxical and inconsistent behavioral changes [1,3]. Our studies in mice identified a phenomenon of hyperlocomotion, an unspecific consequence of chronic stress, which is triggered by a stressful procedure of testing [6,7]. Reduction of the stress impact of testing conditions, e.g., diminishing light intensity, precluded artifacts caused by this phenomenon and let us to determine consistent behavioral phenotype of chronically stressed mice using a variety of tests.

Social behavior as a source of behavioral variability

Even though we used genetically homogeneous mice, we observed pronounced individual variability in animals’ susceptibility to stress-induced anhedonia, which was predicted by subdominant social traits. Stress-induced changes in parameters of sucrose test and other behaviors were depending on percentage of aggressive and non-aggressive individuals in a stressed population, evaluated in baseline conditions [7]. Variability of mouse populations in social behavior presumably underlies significant diversity in behavioral stress response of mice from different batches; in order to ensure a proper stress load, characteristics of stressors had to be adjusted to behavioral parameters of a tested population.

Together, we believe that the findings obtained in our paradigm of stress-induced anhedonia can help to explain and overcome some of the major difficulties in measuring behavior with chronic stress depression models.

References

1. Willner P. (2005). Chronic mild stress (CMS) revisited: consistency and behavioural-neurobiological concordance in the effects of CMS. *Neuropsychobiology* **52**, 90-110.
2. Van den Hove D., Blanco C., Aendekerk B., Desbonnet L., Bruschettini M. Steinbusch H.P., Prickaerts J., Steinbusch H.W. (2005). Prenatal restraint stress and long-term affective consequences. *Dev Neurosci* **27**, 313-320.
3. Nestler E.J., Gould E., Manji H, Buncan M., Duman R., Greshenfeld H., Hen R., Kester S., Ledehendleer I., Meaney M., Robbins T., Winsky L., Zalcman, S. (2002). Preclinical models: Status of basic research in depression. *Biol Psychiatry* **52**, 503-528.
4. Strekalova T., Spanagel R., Bartsch D., Henn F., Gass P. (2004). Stressed-induced anhedonia in mice is associated with deficits in forced swimming and exploration. *Neuropsychopharm* **11**, 2007-2017.
5. Krishnan V., Han M., Graham D., Berton O., Renthal W., Russo S., Laplant Q., Graham A., Lutter M., Lagace D., Ghose S., Reister R, Tannous P., Green T., Neve R., Chakravarty S., Kumar A., Eisch A., Self D., Lee F., Tamminga C., Cooper D., Gershenfeld H., Nestler E.J. (2007). Molecular adaptations underlying susceptibility and resistance to social defeat in brain reward regions. *Cell* **131**, 391-404.
6. Strekalova T., Spanagel R., Dolgov O., Bartsch D. (2005). Stress-induced hyperlocomotion as a confounding factor in anxiety and depression models in mice. *Behav Pharmacol* **16**, 171-180.

7. Strekalova T. (2008). Optimization of the chronic stress depression model in C57 BL/6 mice: evidences for improved validity. *In: Behavioral models in stress research. Nova Sci Publishers NY, US,* 95-139.

Measuring the Impossible: an EU New and Emerging Science and Technologies Initiative Symposium

G.W.A.M. van der Heijden[1], G. Geršak[2], R. Montgomery[3], R. Sefelin[4], and A. Cancedda[5]

[1] Biometris, Wageningen University and Research Centre, Wageningen, the Netherlands gerie.vanderheijden@wur.nl, [2] University of Ljubljana, Ljubljana, Slovenia, [3] National Physics Laboratory, London, UK, [4] CURE, Vienna, Austria, [5] Laboratorio di Scienze della Cittadinanza, Rome, Italy

This symposium will present research funded within the EU "Measuring the Impossible" initiative. The "Measuring the Impossible" initiative is part of the New and Emerging Science and Technologies programs of the EU (http://cordis.europa.eu/nest/).

Measuring the Impossible is about supporting interdisciplinary research and novel investigative methods that could present prospects for advancing the measurement of multidimensional phenomena which are mediated by human interpretation and/or perception – to be able to advance the frontiers of the science of measurement and to respond to future requirements for measuring properties such as comfort, naturalness, perceived quality, feelings, body language and consciousness (ftp://ftp.cordis.europa.eu/pub/nest/docs/1-nest-measuring-290507.pdf).

Many problems in real-life are so complex that they can only be solved using an interdisciplinary approach, including physical metrology (physics, chemistry, biology), neuroscience and psychology. The idea is that solutions require the combined effort of creative researchers with different scientific backgrounds. The emphasis of the initiative therefore lies in the multifaceted character and interdisciplinarity of the research and it aims to promote the creation of new interdisciplinary partnerships between researchers stemming from a wide range of research fields.

Successful cooperation will result in innovative theories and methods for measuring complex human perception and interpretation. The measurement based on novel kinds of technical cognitive systems will help understand human behaviour without neglecting the ethical and gender impacts.

Within the Measuring the Impossible initiative, 15 projects are granted with a total grant from the EU of over €26 M. This symposium will present the progress of seven projects, with emphasis on the techniques and methods that will be used for measuring the complex phenomena.

Symposium contents

MINET – A European network on Measuring the Impossible

M. Gröning

Measuring perceived odor quality

L. Zheng

Focal attention models driven by image statistics

Jan-Mark Geusebroek, Victoria Yanulevskaya, Jan Bernard Marsman, and Frans Cornelissen

Measuring Perception of Naturalness

K.E. Overvliet, S. Soto-Faraco, T.A. Whitaker, L. Johnstone Sorensen, F. McGlone, and G. van der Heijden

Measuring the experience of digital game enjoyment

Wijnand IJsselsteijn

Measuring conscious mental states

M. Overgaard, K. Sandberg, B. Timmermans, and A. Cleeremans

MEMORY: Measuring the relationship between perception of space and time

G. M. Cicchini, P. Binda, D. Burr, and M. C. Morrone

MINET – A European network on Measuring the Impossible

M. Gröning
Department of Psychology/Gösta Ekman Laboratory, Stockholm University, 9 Frescati hagväg, SE-106 91 Stockholm, Sweden, mikael.groning@psychology.su.se

MINET-project fundamentals

The FP6-project "Measuring the Impossible NETwork" (MINET) [1] coordinates the 14 ongoing research-projects within the FP6 activity "New and Emerging Science and Technology" (NEST) action line Pathfinder/Measuring the Impossible (MtI) [2]. The MINET-project is coordinated by Professor Birgitta Berglund at the department of Psychology at Stockholm University and has 22 partners from 11 EU-countries. The project started on February 1st 2007 and will run for 36 months. The total MINET budget is 1.73 million Euros, of which 1.5 million Euros are provided by the EC. The members of the MINET steering committee are: professor Birgitta Berglund, chair, Stockholm University; professor Janko Drnovsek, University of Ljubljana; professor Leslie Pendrill, SP Technical Research Institute of Sweden; Dr. Gerie van der Heijden, Biometris at Wageningen University and Research Centre; professor Giovanni B. Rossi, University of Genova; Ms. Teresa Goodman, National Physical Laboratory.

The Measuring the Impossible

The interdisciplinary nature of the "Measuring the Impossible" area raises many challenges, both in linking related but so far uncorrelated developments in many disparate disciplines/interdisciplines as well as meeting an ever-increasing demand for measurement as a basis for decision making. That is, human interpretation of qualitative information. As proposals for future projects are put forward, it is therefore essential that mechanisms be in place to facilitate science communication and interdisciplinary creativity among the researchers in the Mti-projects, but also in the evolving wider European community.

The MtI initiative focuses on the advancement of interdisciplinary science in soft measurement in any application area. The initiative refers to a variety of different kinds of application areas where *human perception and interpretation* is involved. On the one hand, perception and interpretation are human functions (mental or brain) which are primarily studied in psychology and neuroscience, on the other hand these functions are used for measuring various types of phenomena in a vast number of interdisciplinary application areas.

The measurement of phenomena, which are holistic and multidimensional in nature, is a complex task in itself. In this framework, social aspects of complex phenomena are also involved. It should not be forgotten that human interpretation takes place in the context of social relations, which calls for social science involvement. The research projects under the MtI initiative look to support interdisciplinary research and novel investigative methods that will significantly advance the understanding of human perception and interpretation of complex phenomena.

MINET aims

MINET aims to achieve longterm integration and advancement of the science of measuring complex, holistic quantities and qualities across all relevant areas through Europe. The objectives are to:

1. increase the productivity of MtI projects by interdisciplinary transfer of knowledge on complex measurement, creative interaction and cooperation across application areas;
2. initiate a longterm integration of a broader European community in the advanced science of complex measurement in all relevant fields;
3. facilitate an insightful and wider understanding of the measurement concept across disciplines and its implications in any field of application;
4. improve European scientific competitiveness through the development of competent interdisciplinary expertise and new topics for future MtI projects.

MINET actions

The objectives will be achieved by two main actions:

1. maximizing synergies, boosting creativity and promoting fast progress in the exchange of interdisciplinary knowledge, experience and development of complex measurement through discussion and training: science communication, interdisciplinarity, creativity, terminologies and theories of measurement, and application areas.
2. defining, organizing and managing of a core network of MtI projects, an extended network of the evolving European community including interactive networking with NMIs, and an Internet based Information Center for interactive discussions, dissemination and transfer of scientific knowledge and common activities.

Practical coordination activities are:

a) study-visit program

b) think tanks

c) workshops/seminars

d) training course

Exchange and transfer of knowledge on good practices for complex measurement will positively impact commercial and scientific activities in Europe, e.g., for manufacturing perception delivery on products/buildings/public spaces, and stimulate research projects on new topics.

References

1. www. minet.wordpress.com (2008-06-04)
2. European Commission. Measuring the Impossible – A NEST Pathfinder Initiative. ISBN 92-79-03854-0

SIXTH FRAMEWORK PROGRAMME

Measuring perceived odor quality

L. Zheng

Institute of Environmental Medicine, Karolinska Institute & Department of Psychology, Stockholm University. SE106 91 Stockholm, Sweden, lzg@psychology.su.se

All we know, is that we can discriminate a large numbers of odors, but mechanism to try to describe odor quality is complicated. The perception of odor quality is a homogeneous perception, according to the definition of odor quality, it is "a characteristic of an odor", which we lack a specific language for odor quality. Many studies have been conducted to try to measure the difference between odors. The most efficient way is measure the similarity or dissimilarity between odors. Hedonic factor plays an important role in odor classification and in individual differences found in odor quality studies. Multidimensional scaling has been proved as the most useful method to define odor quality space.

The psychological study of odor quality has been gone through different "trying" steps, the following are some most applied approaches, which were used for measuring the perceived odor quality: (1) *Classification* – group odors by qualitative resemblance [8]. Categorization can be hierarchical which expressed by scheme of odor "primaries"[1]. (2) *Sorting* – sort odors by qualitative resemblance. Subjects sort odors into different numbers of group according to their similarity [11, 12, 15]. (3) *Profiling* – match odor with related word/fixed descriptors [5, 6,13] or with reference odorants [18, 19]. A list of words describing "odor quality" to be attributed to the odor of the inhaled odorants. The attributions are either "objects" and/or "characteristics". (4) *Direct comparison* – judge odors (reference odors and target odors) directly by using free or range of numerical rating of similarity for all pair-wise odors combination [2, 4, 14, 17].

Various statistical measures can be derived from the direct comparison odor/profile data, which were then presented on an odor classification space, each odor represent a point, similar odor stick together form a cluster, the distances (Euclidian distance) between odor points related to the sensory differences between the odors. These measures usually based on two modeling: one is content model, which use principal component analysis (PCA) of similarity and with a vector solution; another one is distance model, multidimensional scaling (MDS) or INDSCAL analysis of 2 way or 3 way [10] data matrices dissimilarity for odor-odor, odor-profile, profile-profile. For sorting data (odorants in groups), the frequencies of the paring of odorants summed over all subjects, those pairings that occur most frequently are considered to have the most similar odors, and least frequency have the least similar odors. This can be used as input for cluster analysis, which can support the MDS solution [11]. The modeling of perceptual structure by the method of multidimensional scaling (MDS) is becoming highly developed and widespread, many applications [2, 4, 7, 11, 14, 15, 17, 18] have been reported Multidimensional scaling (MDS) of similarity judgments provides an ideal method to generate a nonverbal representation of odor quality

The other type of sensory process, which involved in person's perceived odor qualities are: (1) *Odor mixtures*. Independent of type of mixture percept, the mixture quality frequently seems to be intermediate to the quality of its component odors when presented separately [3, 13]. (2) *Adaptation & cross-adaptation*. Repeated exposure will decrease person's ability to discern the differences in odor quality [3]. (3) *Odor intensity*. Concentration changes may also be perceived as odor quality difference [9, 15]. Psychophysical functions, self-adaptation functions, across-adaptation functions and mixture models can measure these effects. The constant of functions used as indicators of "similarity" in odor quality.

Differences in culture, context, social conventions and semantics are known to influence the labeling of odor qualities [15]. Researches on roles of learning in odor perception and on cross-culture differences in odor perception support the same conclusion that experience effects the perceived quality of an odor as well as how much it is like [16].

We use iterative changeover between a top down (perception to molecule) and bottom up approach (molecule to perception) approach to define an odor space. The top down approach is used for determining human ability to differentiate odor quality among single odorous substances, mixtures and complex air mixtures (the question is why are they all different perceived qualities?). The bottom up approach used for selecting reference odor substance (type of molecules or well-controlled mixtures), which together cover the potential odor space. Using this process, we can build up an odor space, which give hierarchical ordinal information on dendrogram, reference odors will define subspaces for repeated testing of new samples of target odorants. Odor space represents an n-dimensional metric space of odor quality.

References

1. Amoore, J.E. (1964). Current status of the steric theory of odor. Ann. NY Acad. Sci., **116**, 457-476.
2. Berglund, B., Berglund, U., Engen, T., & Ekman, G. (1973). Multidimensional analysis of twenty-one odors. *Scand J. Psychol*.,**14**, 131-137.
3. Berglund, B. (1974). Quantitative and qualitative analysis of industrial odors with human observers. Ann. NY Acad. Sci., **237**, 35-51.
4. Davis, R. G. (1979). Olfactory perceptual space models compared by quantitative methods. *Chemical Senses and Flavour*, **4(1)**, 21-33.
5. Dravnieks, A. (1982). Odor quality: semantically generated multidimensional profiles are stable. *Science*, 218, 799-801.
6. Dravnieks, A., Bock, F. C., Powers, J. J., Tibbetts, M., & Ford, M. (1978). Comparison of odors directly and through profiling. *Chemical Senses and Flavour*, **3(2)**, 191-225.
7. Gregson, R. A. M., & Mitchell, M. J. (1974). Odor quality similarity scaling and odor word profile matching. *Chemical Senses and Flavor*, **1**, 95-101.
8. Harper, R., Bate Smith, E.C. & Land, D.G. (1968). *Odour Description and Odour Classification*. American Elsevier, New York.
9. Gross-Isseroff, R., & Lancet, D. (1988). Concentration-dependent changes of perceived odor qulity. *Chem.Senses*, **13**, 191-204.
10. Kruskal, J.B. (1964). Nonmetric multidimensional scaling. A numerical method. *Psychometrika*, **29**, 1-27, 115-129.
11. Lawless, H. T. (1989). Exploration of fragrance categories and ambiguous odors using multidimensional scaling and cluster analysis. *Chemical Senses*, **14(3)**, 349 -360.
12. MacRae, A. W., Rawcliffe, T., Howgate, P., & Geelhoed, E. N. (1992). Patterns of odour similarity among carbonyls and their mixture. *Chemical Senses*, **17(2),** 119 -125.
13. Moskowitz, H.R., & Barbe, C.D. (1977). Profiling of odors components and their mixtures. *Sens.Process*., **1**, 212-226.

14. Moskowitz, H. R., & Gerbers, C. L. (1974). Part I. Functional properties of the olfactory system: psychophysics: dimensional salience of odors. *Ann N Y Acad. Sci*., 237, 1-16.
15. Stevens, D. A., & O'Connell, R. J. (1996). Semantic-free scaling of odor quality. *Physiology & Behavior*, 60, 211-215.
16. Stevenson, R. J., & Boakes, R. A. (2003). A mnemonic theory of odor perception. Psychological *Review*, 110, 340-364.
17. Woskow, M.H. (1968). Multidimensional scaling of odors. In Tanyolac, N. (ed.), *Theories of Odor and Odor Measurement*. Robert College, Istanbul, pp. 147-191.
18. Wright, R.H., & Michels, K.M. (1964). Evaluation of infrared relations to odor by a standards similarity method. *Ann. NY Acad. Sci*., 116, 535-551.
19. Yoshida, M. (1975). Psychometric classification of odors. *Chem. Senses Flav*., 1, 443-464.

Focal attention models driven by image statistics

Jan-Mark Geusebroek[1], Victoria Yanulevskaya[1], Jan Bernard Marsman[2], and Frans Cornelissen[2]
[1]*Informatics Institute, University of Amsterdam, The Netherlands, mark@science.uva.nl*
[2]*School of Behavioural and Cognitive Neurosciences, University Medical Centre Groningen, The Netherlands*

The visual system evolved to provide fast and adequate, and not necessarily the best, estimates of physical properties. One way in which the visual analysis has been speeded up is by analyzing only part of the scene in detail, with our foveae. The trade-off is that we must move our eyes to areas of interest for detailed analysis. Deciding where to move requires a fast selection process, ideally based on a moment-to-moment representation of potentially "important" regions in the natural scene. We propose the visual system does so by determining areas that deviate in a statistical sense from the rest of the visual scene. This information constitutes a statistical salience map that will drive gaze and attentional shifts.

The concept of the salience map is central in the study of attention and eye movements. A number of neurally plausible computational schemes have been proposed for the creation of salience maps, most notably Li's [1] work on V1, on which our hypothesis is based, but see also [3,4]. The important new aspect that this study brings is that, unlike previous imaging and most behavioral studies, we will start from a theoretical understanding of the statistical characteristics of natural images [5,6]. Statistical analysis of natural scenes and measurement-theoretical assumptions can tell us what information can be retrieved from the visual environment, with initial experiments shown in [7].

In computer vision, the selection of salient regions in an image is a well studied topic, for a review see [8], being the first step in many computer vision algorithms. Combining the insights from eye tracking studies with theoretical models of salient regions detection, we aim to arrive at better prediction of regional image saliency. Natural image statistics provide a means to quantize saliency, by comparing the common visual structures of the world around us, in contrast to salient visual structures which stand out by their sparse occurrence in the visual field. Particularly, we consider the integrated Weibull distribution as a parameterized model, which provides a good fit to the statistics of natural images [5]. We show how distinct regimes of the integrated Weibull distribution leads to various local saliency mechanisms. With model selection techniques from information theory, we can determine the probability for every distinct regime, to explain the statistical properties of local image content. These lead to different mechanisms for saliency determination, see Figure 1 for an example.

Measurement of saliency from natural image statistics

The contrasts in a large range of natural scenes conform to a Weibull type of distribution [5]. The integrated Weibull distribution includes the class of symmetric exponential distributions. Its shape indicates different regimes of the distribution (see Figure 2), thereby roughly categorizing image content [9]. When the contrasts in the image constitute a power-law, this indicates a contrasting foreground object against a background. When more and more clutter occurs, the shape progresses from power-law through exponential to become Gaussian for high frequent textures. Images with a regular pattern can not be described well with the integrated Weibull distribution. Note that parameters can be estimated with Maximum Likelihood Estimation (MLE) technique. It is possible to distinguish tree types of images according to the behavior of integrated Weibull distribution, or conclude that integrated Weibull distribution does not describe data well. For the first case, we use Akaike's information criterion (AIC) for appropriate model selection [10]. AIC estimates expected Kullback-Leibler information, based on the log-likelihood function at its maximum point. Hence, we do not need to assume that the "true model" is in the set of candidates. Regarding the latter case, we use Anderson-Darling goodness of fit test at a 5% confidence level. The Anderson-Darling test is a generalization of the Kolmogorov-Smirnov test, which is more sensitive to deviations in the tails of the distribution. This is a relevant characteristic of the test, as in our case the tails capture the strong, hence important contrasts present in the scene. We applied the approach to 50 by 50 patches from a hundred natural images (1 megapixels) taken from the National Geographic website. Of the patches, 72% were Weibull distributed of which 22% power-law, 42% exponential, and 8% Gaussian, respectively. The remaining 28% of the patches was rejected by the Anderson-Darling test, and constitutes a mixture of Weibulls [11], or regular patterns [6], experimentally evaluated in Figure 3. Each of these distributions indicates a different mechanism for regional saliency.

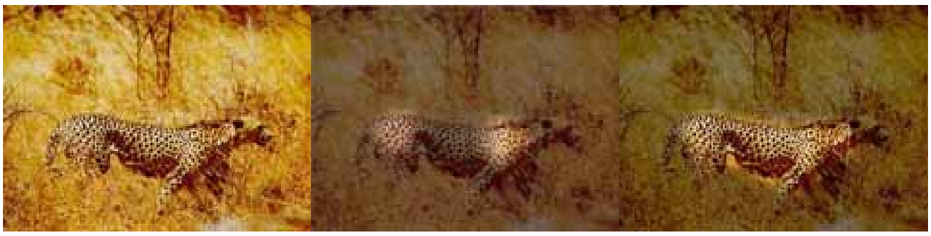


Figure 1. *An example of the use of the Weibull distribution in the determination of salient regions. Middle image shows example fixation points, sparsely sampling the object (original left). On the right, the region which deviates from the common statistics in the scene is highlighted, as determined by testing goodness-of-fit over an extended region.*

References

1. Li, Z. (2002). A salience map in primary visual cortex. *TINS*, **6**, 9-16.
2. Itti L., Koch C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Res*. **40**, 1489-1506.
3. Deco G., Rolls E.T. (2004). A neurodynamical cortical model of visual attention and invariant object recognition. *Vision Res*. **44**, 621-642.
4. Geusebroek J.M., Smeulders A.W.M. (2003). Fragmentation in the vision of scenes. *Int. Conf. Comput. Vision*, 130-135.
5. Geusebroek J.M., Smeulders A.W.M. (2005). A six-stimulus theory for stochastic texture. *Int. J. Comput. Vision*, **62**, 7-16.
6. Marsman J.B.C., Yanulevskaya V., Cornelissen F.W., Geusebroek J.M. (2007). Spatial statistics of natural images predict gaze direction, Perception **36** E*CVP Abstract Supplement.*
7. Mikolajczyk K., Schmid C. (2004). Scale & affine invariant interest point detectors. *Int. J. Comput. Vision* **60**, 63–86.
8. van Gemert J.C., Geusebroek J.M., Veenman C.J., Snoek C.G.M., Smeulders A.W.M. (2006). Robust scene categorization by learning image statistics in context, *Proc. CVPR Workshop on SLAM.*
9. Burnham K.P., Anderson D.R. (2002). Model selection and multimodel inference: A practical information-theoretic approach. Springer.
10. Burghouts G.J., Smeulders A.W.M., Geusebroek J.M. (2007). The distribution family of similarity distances. *Proc. Advances Neural Inf. Process. Syst*., 2007.

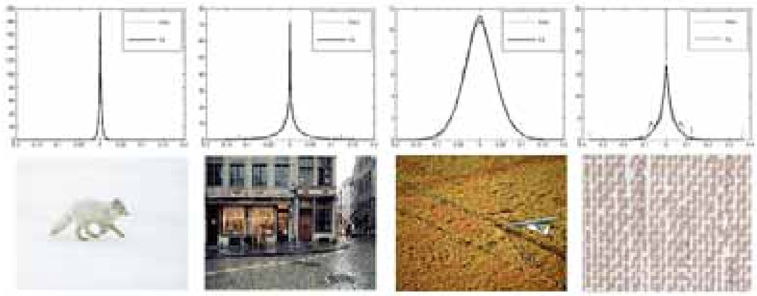


Figure 2. *The integral form of the Weibull distribution fits the statistics of contrasts in natural images very well. The distribution is characterized by a width and an exponent parameter. Its shape is exponential, where the exponent parameter determines shape from power-law like (most left) for single contrasting objects against background; to double exponential (left) for more fragmented scenes; to Gaussian (right) for high frequent textures. If the scene contains a close to regular pattern, the integrated Weibull does not fit (most right). Hence, fitting a Weibull distribution measures various regimes of natural image statistics.*

Figure 3. *Results from an eyetracker experiment testing saccade behaviour for non-Weibull versus Weibull textures. Levels indicate the probability of making a saccade towards a non-Weibull texture out of 8 circular arranged textures. Red denotes results from 15 subjects for the CuRet dataset (9 non-Weibull, 52 Weibull); Blue indicates data from 9 subjects for the Alot dataset (34 non-Weibull, 216 Weibull).*

Measuring Perception of Naturalness

K.E. Overvliet[1], S. Soto-Faraco[1], T.A. Whitaker[2], L. Johnstone Sorensen[3], F. McGlone[4], and G. van der Heijden[5]
[1]*Parc Científic de Barcelona, Universitat de Barcelona, Spain, krista.overvliet@gmail.com*
[2]*School of Psychology and Institute of Neuroscience, Trinity College Dublin, Ireland, whitaket@tcd.ie*
[3]*National Physical Laboratory (NPL), London, UK, linda.sorensen@npl.co.uk*
[4]*Unilever Research Port Sunlight Laboratory, UK, francis.mcglone@unilever.com*
[5]*Biometris / Plant Research International, Wageningen University and Research Centre, The Netherlands, gerie.vanderheijden@wur.nl*

Measuring Naturalness

Naturalness is often regarded as a highly valued property of a material. Most of us, for example, would accept to pay more for a natural silk dress or wooden floor than for functionally equivalent but "fake" versions, meeting the same standards. What physical properties are underlying naturalness perception? And how do people know the difference? The MONAT (Measurement of Naturalness) project aims to understand how the perception of naturalness is formed.

Each material has specific physical properties that differentiate them from other materials. These properties are picked up by one or several of our sensory systems and bound into a unitary percept. For example, properties like colour and glossiness are perceived by our visual system, whereas temperature and roughness are encoded by the tactile system. At some stage of processing these, initially separate, perceptual properties are combined to form a unitary percept of the material upon which any judgments about its naturalness will be based.

Using a multidisciplinary approach, we seek to better understand the relationship between the physical measurements of material properties and the perceptual as well as neural bases of naturalness judgements by humans. In this paper we will focus on the subjective measurements of human naturalness judgements, and the contributions of vision and haptics to this perceptual decision.

Psychophysical Measurements of Naturalness

Psychophysics is a sub-discipline of psychological measurements which aims to unravel the relationship between objective characteristics of physical stimuli and their subjective interpretation by a human observer. A common problem in psychophysics is establishing a link between the (subjective) perception being measured and the (objective) physical parameters being manipulated. Because naturalness is, in itself, a rather elusive concept, psychophysical measurement is extremely difficult. To tackle similar problems in the past, researchers have made use of converging measures from different psychophysical paradigms. The rationale is that if several measuring methods provide consistent subjective estimates, one can conclude that they are measuring the same perceptual property, and that such a property has a psychological status. Here, we used an approach based on direct measurements of naturalness: labelled category scaling, free modulus magnitude estimation and a two- alternative forced choice (2AFC) task. In this paper we investigate whether these three ways of measuring naturalness judgments are actually measuring the same underlying property (i.e., naturalness), and what the contributions of vision and haptics are to the perception of this property.

In all three methods we used the same experimental setup. Participants were seated behind a table in such a way that the angle between the thorax and the table was approximately 90°. The height of the chair was adjusted so that they could comfortably reach the stimuli. The stimuli were 30 different samples of real wood, laminate, veneer, vinyl and photocopy, all being oak or imitations thereof. They were mounted behind a window of 8 by 8 cm in the top of a square, gray plastic box, so that the participant was only able to see and/or feel the top surface of the material sample. The samples were placed one by one in a standard daylight tent for 3 seconds. In the visual exploration conditions, the participant was only allowed to look at the stimuli. In the tactile exploration condition a curtain was placed between the participant and the stimulus to prevent them from seeing the sample. In this condition, they were asked to perform 3 one-second circular movements with the index finger of their dominant hand on the surface of the stimulus. In the visuo-tactile exploration condition the participant was asked to use both these exploration strategies simultaneously. In all cases, after three seconds exploration time, observers had to verbally state their judgement (see measurement methods, below). The measurement methods were counter-balanced and the stimuli were presented in random order for each observer and condition.

Labelled Category Scaling

In labelled category scaling the participant is required to label each stimulus according to a pre-defined scale with labelled categories. Rozin [1] investigated the perception of naturalness in the domain of foods using a labelled category scale, so we decided to use a Spanish version of his scale. In table 1 we show the original (English) version.

Table 1. *Labbeled scale of judgement of naturalness used in the experiment (the English original [1])*

0	Not natural at all
1	Very slightly natural
2	Slightly natural
3	Moderately natural
4	Very natural
5	Extremely natural
6	Completely natural

Free Modulus Magnitude Estimation

Free-modulus magnitude estimation is based on the magnitude estimation method as described by Stevens [2]. Participants were asked to assign any (arbitrary) numerical value to the first stimulus, and then assign numbers to the following stimuli accordingly, trying to capture the ordering and subjective distance between stimuli for the judged property (naturalness, in this case). They were allowed to use any number they liked. Unlike in the labelled category scaling, we explicitly asked the participant to compare the different stimuli to each other. The numerical scores the participants had given were re-scaled linearly (between 0 – 6) in order to make this measurements readily comparable between participants and to the other methods.

Binary Decision Task

In the binary decision task, the participants were asked to judge whether they think a stimulus is natural or not. They simply responded with a "yes" or a "no". We averaged the decisions of all the participants and rescaled the results (from 0 to 6) to make the scores comparable to the other methods.

Results and Discussion

The initial results indicate that the measurements are highly correlated, thus suggesting that they measure a common underlying construct. Further analysis and experiments might be needed to investigate whether the concept that we are measuring was actually "naturalness" and not for example one of the characteristics of the wood (e.g. roughness). We will therefore link the measured physical characteristics to the three methods of psychological judgements. Moreover, we will measure naturalness in other materials using the same methods.

Another important finding from the present study is that the visual and haptic estimates of naturalness are moderately correlated, suggesting that their contributions are relatively independent. The combination of visual and haptic estimates, when both are available, seems to amount to a weighted average of the two individual estimates.

References

1. Rozin, P.(2005). The Meaning of "Natural" Process More Important Than Content.*Psychological Science*. **16(8)**, 652-658.
2. Stevens, S.S. (1975). Psychophysics: Introduction to its Perceptual, Neural and Social Prospects, ed. G. Stevens. New York: Wiley

Measuring the Experience of Digital Game Enjoyment

Wijnand IJsselsteijn[1], Wouter van den Hoogen[1], Christoph Klimmt[2], Yvonne de Kort[1], Craig Lindley[3], Klaus Mathiak[4], Karolien Poels[1], Niklas Ravaja[5], Marko Turpeinen[6], and Peter Vorderer[7]

[1] *Eindhoven University of Technology, Eindhoven, The Netherlands, w.a.ijsselsteijn.tue.nl*

[2] *Hochschule für Musik und Theater, Hannover, Germany*

[3] *Blekinge Institute of Technology, Karlshamn, Sweden*

[4] *RWTH Aachen University, Aachen, Germany*

[5] *CKIR, Helsinki School of Economics, Helsinki, Finland*

[6] *Helsinki Institute for Information Technology, Helsinki, Finland*

[7] *CAMeRA, Free University of Amsterdam, Amsterdam, The Netherlands*

The area of digital games constitutes a tremendously varied set of applications, with a wide range of associated player experiences, defying a one-size-fits-all approach to its conceptualization and measurement. One of the main challenges facing the gaming research community is a lack of a coherent and fine-grained set of methods and tools that enable the measurement of entertainment experiences in a sensitive, reliable and valid manner. Much like the six wise (but blind) men touching the elephant, no single methodological perspective can be said to provide a comprehensive understanding of digital gaming. Following this insight, the FUGA project, funded under the EU FP6 NEST 'Measuring the Impossible' initiative, takes a critical view towards the exclusive reliance on any one single indicator for measuring player experience. In the FUGA project, we explicitly strive towards a multi-method, multi-measure approach whereby we anchor and cross-validate various measures (e.g., self-report, psychophysiological, behavioural, neural) via their simultaneous application to a certain standardized set of games, and correlating the results thus obtained. Assessing the basic psychometric properties (sensitivity, reliability, validity) of all measures developed in the project is one of the defining characteristics of FUGA.

As a significant first step, we have developed and validated the Game Experience Questionnaire (GEQ), which reliably distinguishes between seven different dimensions of player experience: Sensory and Imaginative Immersion, Tension, Competence, Flow, Negative Affect, Positive Affect, and Challenge [4][6][7]. The GEQ is freely available in Dutch, English, German, Swedish and Finnish. Translations in other languages, including French, are in progress. At the moment, the GEQ is being applied in a number of experimental research studies within the FUGA project, and results testify to the measure's test-retest reliability and construct validity. Further to self-report measures, the FUGA project is investigating a number of implicit and objective measures of player experience. Overt (e.g., facial expressions) and covert (e.g., pressure exerted on an interaction device) expressions of behaviour are being investigated for their potential to validly and reliably tap in on certain dimensions of player experience, such as boredom, flow and frustration. First results are encouraging, showing a positive correlation between pressure exerted on the left ('fire') button of the computer mouse, the amount of bodily movement a player exhibits, and several player experiences, including frustration [3]. Psychophysiological recordings (e.g., EEG, facial EMG, EDA) show particular promise in distinguishing player emotions along the dimensions of arousal (exciting vs. dull) and valence (positive vs. negative emotions) [8]. In addition, various studies have been carried out that investigate the effects of collaboration vs. competition [9][10], something that differentially affects the experience of male and female players. Moreover, there are initial indications that results obtained in the laboratory do not differ significantly from results obtained in the field, which bodes well for the external validity of psychophysiological measures [9]. An implicit association test [1] was successfully developed and applied to detect identification processes in games, exploring theory-consistent implicit cognitions that link game character attributes to the player's self perception [2]. Furthermore, functional Magnetic Resonance Imaging (fMRI) has been applied to investigate the neural correlates associated with particular player experiences [5], while in the process developing innovative procedures for executing a think-aloud protocol while being in an fMRI machine.

We believe that a large range of measures, from reflective (subjectively controllable) to fully reflexive (uncontrollable) responses, enables a fuller characterization of the game experience than any single isolated measure, thus sensitizing us to the rich gamut of experiences associated with digital games. Moreover, limitations particular to one measure may be overcome or compensated by using corroborating evidence emerging from another measure. The combination of multiple measurement modalities can thus reduce uncertainty associated with measuring a single modality, resulting in increased validity, robustness and wider applicability of the total set of measures.

Finally, to the extent that a number of these measures allow for *real-time* measurement of player experiences, they can potentially be applied as continuous input data to a game engine, allowing the game's AI to appropriately adjust to the player's experiential (affective) state at any point during gameplay. Such a closed loop between player experience and game dynamics could help create an exciting new genre of emotionally adaptive digital games, and would allow for highly personalized optimization of such games. Under FUGA, several prototypes of such emotionally adaptive games have been created.

References

1. Greenwald, A. G., McGhee, D. E., & Schwartz, J.L.K. (1998). Measuring individual differences in implicit cognition: The Implicit Association Test. *Journal of Personality and Social Psychology*, **74** , 1464-1480.
2. Hefner, D., Klimmt, C., & Vorderer, P. (2007). Identification with the player character as determinant of video game enjoyment. *Proceedings of ICEC 2007, 6th International Conference on Entertainment Computing*, 39-48.
3. van den Hoogen, W., IJsselsteijn, W.A., de Kort, Y.A.W., Poels, K. (2008). Towards real-time behavioral indicators of player experiences: Pressure patterns and postural responses. *Proceedings of Measuring Behavior 2008 (Maastricht, The Netherlands, 26-29 August 2008).*
4. IJsselsteijn, W.A., de Kort, Y.A.W., & Poels, K. (n.d.). The Game Experience Questionnaire: Development of a self-report measure to assess player experiences of digital games. Manuscript submitted for publication.
5. Klasen, M., Weber, R., Mathiak, K.A., Zvyagintsev, M., & Mathiak, K. (2008). Think aloud-method during fMRI to

determine neuronal correlates of subjective experience of video game playing. *Proceedings of Measuring Behavior 2008 (Maastricht, The Netherlands, 26-29 August 2008).*

6. de Kort, Y. A. W., IJsselsteijn, W. A., & Poels, K. (2007). Digital games as social presence technology: Development of the Social Presence in Gaming Questionnaire (SPGQ). *Proceedings of PRESENCE 2007, The 10th International Workshop on Presence (Barcelona, Spain, 25-27 October 2007),* 195-203.
7. Poels, K., de Kort, Y.A.W., & IJsselsteijn, W.A. (2007). "It is always great fun!" Exploring dimensions of digital game experience using focus group methodology. *Proceedings of the 2007 Conference on Future Play (Toronto, Canada, 14-18 November 2007),* 83-89.
8. Ravaja, N. & Kavikangas, M. (2008). Psychophysiology of digital game playing: The relationship of self-reported emotions with phasic physiological responses. *Proceedings of Measuring Behavior 2008 (Maastricht, The Netherlands, 26-29 August 2008).*
9. Ravaja, N. & Kavikangas, M. (2008). Psychophysiology of digital game playing: Effects of competition versus collaboration in the laboratory and in real life. *Paper presented at ICA 2008 (Montreal, Canada, 22-26 May 2008).*
10. Ravaja, N., Turpeinen, M., Saari, T., Puttonen, S., & Keltikangas-Järvinen, L. (2008). The psychophysiology of James Bond: Phasic emotional responses to violent video game events. *Emotion*, **8**, 114-120.

Measuring conscious mental states

M. Overgaard[1], K. Sandberg[1], B. Timmermans[2], and A. Cleeremans[3]
[1]CNRU, Hammel Neurorehabilitation and Research Unit, Aarhus University Hospital, Hammel, Denmark, neumov@sc.aaa.dk
[2] CCC, Universite Libre de Bruxelles, Brussels, Belgium
[3]CSRU, Universite Libre de Bruxelles, Brussels, Belgium

How can we measure whether a particular mental event is consciously experienced or not? Such measurements provide the essential data on which the current and future science of consciousness depends, yet there is little consensus about what would suffice as a measure of consciousness. Having dependable measures of consciousness is essential not just to arrive at empirically based theories of consciousness, but for any other discipline interested in experimental or clinical studies of conscious subjects.

The NEST-PATHFINDER project *MindBridge* studies the topic of human consciousness from a multidisciplinary perspective. Human consciousness can be defined as the subjective experience of mental states such as perceptions, judgments, thoughts, intentions to act, feelings or desires

[1]. These experiences can only be described meaningfully from a subjective, phenomenal first-person account [2]. On the other hand, cognitive neurosciences explore brain events associated with cognition from an objective third-person account [3]. It is the mission of *MindBridge* to develop strategies and concrete methodologies that can bridge this gap between subjective experience and objective observation of neural phenomena.

As seen in figure 1 [4] illustrates how any correlation between conscious states and brain states are logically dependent on the strength of the relations between the actual states and the methods that claim to measure them. The relation between the conscious state and the "behavioural indication" in the figure is essentially a question of the relation between the experience, as it is for the subject, and some report issued by the subject.

We present a series of experiments, comparing current behavioral methods in consciousness studies in order to arrive at specific recommendations for how to measure consciousness. The main finding for use in future experiment is that subjective reports seem to be more valid when using a gradual rather than a dichotomous scale. Although this is still speculative, the important factor may be that subjects are able to relate specific conscious states, as they are experienced, to specific reported categories. In this respect, the applied 4-point scales seem effective [5]. It is not advised to continue using the dichotomous reports which are often seen in the experimental literature on consciousness.

The results, furthermore, suggest that consciousness is gradual rather than dichotomous in its nature.

Touching briefly on other aspects of research conducted by the *MindBridge* consortium, we will show how the development of methods to study consciousness has important bearings on the way clinical work in cognitive neurorehabilitation is carried out. We will present experimental data of studies on conscious perception in a blindsight and 7 neglect patients, arguing that a sensible handle to measure conscious states is vital in the assessment of brain injured patients and in the decision, which kinds of cognitive training should be applied.

References

1. Hameroff, S. (2001): Consciousness, the brain, and spacetime geometry, *Annals of the New York Academy of Sciences,* **929,** 74-104
2. Overgaard, M. (2004): Confounding factors in contrastive analysis , *Synthese,* **141,** 217-231
3. Seth, A. Dienes, Z., Cleeremans, A. Overgaard, M. & Pessoa, L. (in press): Measuring consciousness: Relating behavioural and neurophysiological approaches, *Trends in Cognitive Sciences*
4. Overgaard, M. (2006): Introspection in science, *Consciousness and Cognition,* **15**, 629-633
5. Ramsøy, T.Z. & Overgaard, M. (2004): Introspection and subliminal perception, *Phenomenology and the Cognitive Sciences,* **3**, 1-23

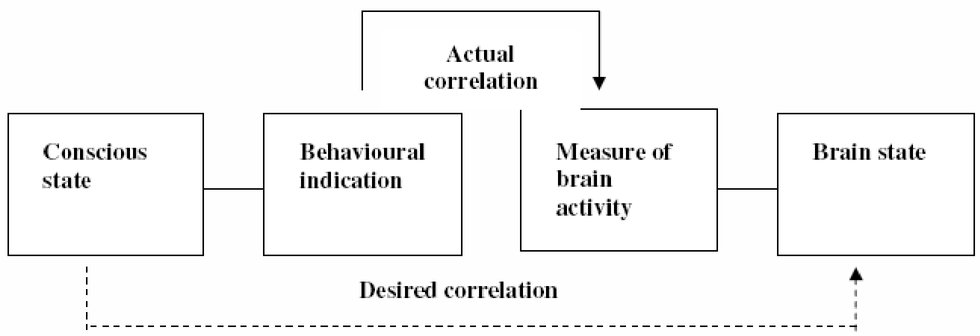


Figure 1. *The relation between conscious and brain states in the search for neural correlates of consciousness*

MEMORY: Measuring the relationship between perception of space and time

G. M. Cicchini[1], P. Binda[1], D. Burr[2], and M. C. Morrone[1]
[1]*Dipartimento di Psicologia, Università Vita-Salute, San Raffaele, Milan, Italy, g.cicchini.hsr.it*
[2]*Dipartimento di Psicologia, Università degli Studi di Firenze, Florence, Italy*

Time and space are normally considered to be independent, and are studied typically separately by neuroscientists. However, evidence is beginning to emerge that they are in fact strongly interconnected in the brain [1, 2]. Experiments in our laboratory, using standard techniques of adaptation, show that neural units sensing the time of events are spatially selective [3]. In these experiments observers adapted to a grating drifting within a 12° diameter patch while fixating on a spot to its lower left; once the adaptation period elapsed, they changed their gaze to a target that appeared 15° right of fixation. Eight hundred ms after the display of the target, a test grating was presented for 600 ms in one of three randomly chosen positions: the same retinotopic position, the same spatiotopic position or in one that was neither (control condition). Perceived duration in the control and the retinotopic locations were near 600 ms, the actual duration of the test. On the other hand events in the spatiotopic location were affected by adaptation and were perceived on average as 456 ms. These results imply that there are multiple timing mechanisms for visual events; these units are selective in real-world coordinates, taking into account the position of gaze.

Eye movements themselves also affect the perception of time. At the time of saccades, temporal intervals marked by two brief visual stimuli are strongly underestimated. A test interval of 100 ms is judged veridically as long as it is presented far from an eye movement [2]. However, if the test interval is presented perisaccadically, its duration is compressed down to 50ms. The effect occurs for intervals presented from 150 ms before a saccade to 150 ms after the onset of the eye movement. The temporal compression parallels closely perisaccadic mislocalization, a phenomenon which is thought to be the by-product of the remapping of retinal information which compensates for gaze shifts. Not only: under some conditions the perceived temporal order of two bars flashed just before a saccade can even be reversed [2, 4]. The moment at which temporal order inversion is most prominent coincides with the beginning of spatial remapping. These results suggest a strong correlation between the mechanisms mediating perception of space and perception of time [2, 5].

More recent experiments show that it is not actually necessary to move the eyes for the perception of time to be compressed: it suffices to transfer attention from one object to another [6]. Notably, also in this case the effects are spatially selective. In this paradigm subjects were asked to perform concurrently a visual discrimination task (primary task) and a duration discrimination task (secondary task). We tested various positions both tasks. In one condition the two bars marking the duration, were flashed in two separate locations, in a second condition the two bars were flashed in the same spatial location. We found that shifting attention to perform the primary task can induce a strong compression of temporal intervals. However, this compression of perceived time occurs only if the subject needs to integrate temporal information signalled in two separate locations. Intervals marked in the same spatial position were not subject to such a strong temporal compression and were not affected by attentional deprivation. This indicates that location where the events take place is crucial in determining perceived duration.

Taken together, these experiments show that neural representations of space and time are strongly inter-dependent.

Acknowledgements

This research has been funded by the EU-NEST project MEMORY, FP6 EU-NEST 43236

References

1. Johnston, A., D.H. Arnold, and S. Nishida. (2006) Spatially localized distortions of event time. *Curr Biol.*. **16(5)**: p. 472-9.
2. Morrone, M.C., J. Ross, and D. Burr. (2005) Saccadic eye movements cause compression of time as well as space. *Nat Neurosci.*. **8(7)**: p. 950-4.
3. Burr, D., A. Tozzi, and M.C. Morrone. (2007) *Neural mechanisms for timing visual events are spatially selective in real-world coordinates.* Nat Neurosci. **10(4)**: p. 423-5.
4. Binda, P., D.C. Burr, and C.M. Morrone. (20080) Spatio-temporal distortions of visual perception during saccades. in *Abstracts of the 30th European Conference in Visual Perception. Arezzo (27th - 31st August 2008)* Perception 36s: p. 112.
5. Morrone, M.C., J. Ross, and D.C. Burr. (2008) Keeping vision stable: rapid updating of spatiotopic receptive fields may cause relativistic-like effects., in *Space and time in perception and action*, R. Nijhawan, Editor. Cambridge University Press: Cambridge.
6. Cicchini, G.M. and C.M. Morrone. (2008) Shifts in spatial attention affect the perceived duration of events. in *Abstracts of the 30th European Conference in Visual Perception. Arezzo (27th - 31st August 2008)* Perception 36s: p. 113.

Measuring Player Experiences in Digital Games: Behavior, Psychophysiology, and Brain-Based Indicators Symposium

Wouter van den Hoogen
Game Experience Lab, Human Technology Interaction Group, Eindhoven University of Technology, Eindhoven, The Netherlands, W.M.v.d.Hoogen@tm.tue.nl

The game industry has been booming for many years. A wide range of games and gaming devices have been developed targeted to different age and gender groups and to gamers with various play styles. As such, playing digital games has become a highly popular leisure activity. Over recent years, academic research on digital gaming has however been growing in interest. Part of the research is focused on how to measure what people feel and experience when they engage in playing digital games. To date however, research focussed on establishing valid behavioural indicators of people's experiences with digital games have been conducted on a modest scale. Being able to validly and reliably measure these experiences is a prerequisite for sound research and theory building in this domain.

The aim of the symposium is to establish ways of modelling people's experiences with interactive media in a real-time fashion. How can these experiences be measured in a valid way beyond retrospective self-reports, and are there ways of measuring people's experiences with game-play real-time? These questions are central to the presentations in the symposium.

Continuous measures, that are collected as the experience is unfolding, have a number of advantages over retrospective evaluations. For one, they do not rely on memory. As experience is very likely to change dynamically during play, continuous measures provide a richer and presumably more valid evaluations pattern compared to retrospective judgments. Secondly, with the exception of think-aloud protocols, continuous measures generally do not require introspection, but rather make use of objective indicators of experiences and thus nicely complement the subjective self-report based measures. Thirdly, being able to register experiences during game play, without having to interfere with the game play or without having to interrupt the gamer, builds opportunities for directly studying which specific events or episodes in a game evoke the most intense or engaging experiences. To date, however, validated tools that enable researchers to evaluate game experience in a continuous way are rather scarce. While there is a body of research aimed at measuring behaviour in persons to person situations, knowledge about people's behaviour when they play digital games is only beginning to emerge.

In this symposium five papers are presented that advance the state-of-the-art in continuous measurement of digital game experiences, including real-time behaviour tracking, observational coding, psycho-physiological measures and approaches from brain computer interaction. The papers will address a number of questions regarding continuous measures, such as: what indicators of players' experience can be recorded during game play; how are they related to (components of) game experience; what methods can be employed to analyze game experience per se, and in relation to specific game events or episodes.

Symposium contents

A physiological approach for continuously modeling user emotion in interactive play environments
R.L. Mandryk

Think Aloud-method during fMRI to determine neuronal correlates of subjective experience of video game playing
Martin Klasen, René Weber, Krystyna A. Mathiak, Mikhail Zvyagintsev, and Klaus Mathiak

Multi-modal Behavioral Cues from Bodily Interaction in Ambient Entertainment Applications
Anton Nijholt, Betsy van Dijk, and Dirk Heylen

Psychophysiology of digital game playing: The relationship of self-reported emotions with phasic physiological responses
N. Ravaja and J.M. Kivikangas

Towards real-time behavioral indicators of player experiences: Pressure patterns and postural responses
Wouter van den Hoogen, Wijnand IJsselsteijn Yvonne de Kort, and Karolien Poels

A physiological approach for continuously modeling user emotion in interactive play environments

R.L. Mandryk
Department of Computer Science, University of Saskatchewan, Saskatoon, Canada, regan@cs.usask.ca

Computer games have grown during recent years into a popular entertainment form with a wide variety of game types and a large consumer group spread across the world. As researchers develop novel play environments, computer and console game markets continue to grow rapidly, outperforming the film industry in terms of total revenues in many regions [6]. Although gaming technology has continued to evolve, researchers and computer game developers suffer from a lack of effective evaluation methods. My research interest is in how to quantify emotional experience when engaged with play technologies, by developing an evaluation methodology for entertainment environments that is as robust as methods for evaluating productivity. This paper summarizes a series of experiments used to design this new evaluative methodology.

Motivation

The development of evaluation methodologies in human-computer interaction research (HCI) has been rooted in the sciences of Psychology, Human Factors, Engineering and Computer Science [5]. The idea of emotion, which is equally important to design [5], is still not well understood, especially when the primary goals are to challenge and entertain a user. Traditional measures of user behaviour in productivity environments, such as task performance, are not applicable to play environments since we are not interested in measuring performance; we are interested in measuring what kind of emotional experience is provided by the play technology and environment, regardless of performance [6]. Although traditional usability measures may still be relevant, they are subordinate to the emotional experiences resulting from interaction with the play technology and with other players. An equally important issue for measuring user experience in game environments is that a successful play experience is determined by the process of playing, not the outcome of playing [6]. Any methodological approaches should inform designers of user behaviour both within and after the play experience.

Experiments

Along with several co-authors, I conducted a series of experiments to develop a new method for continuously modeling user emotional state in interactive play environments based on a user's physiological responses. The goal was to develop an evaluation methodology for entertainment environments that: captures usability and playability through metrics relevant to ludic experience; accounts for user emotion; is objective and quantitative; and has a high evaluative bandwidth (continuous measurement). Our approach was to collect physiological data as a direct indication of user experience. We explored the following physiological measures: galvanic skin response (GSR), heart rate (HR), electromyography (EMG) of the face (jaw, forehead, cheek), and respiration.

In an initial experiment we explored how physiological signals respond to interaction with play technologies. We collected a variety of physiological measures while observing participants playing a computer game in four difficulty conditions, providing a basis for experimental exploration of this domain. Collecting and analysing physiological data requires a controlled approach that is hard to balance with the need for ecological validity when measuring behaviour within gaming systems. Guidelines for conducting research in this domain based on results from our initial experiments can be found in [3].

In a second experiment we investigated how physiological signals differ between play conditions, and how physiological signals co-vary with subjective reports [1]. We observed different physiological responses when users played a computer game against a co-located friend versus a computer. When normalized, the physiological results mirrored subjective reports. By integrating the methods in [3], we showed that physiological measures can be used to objectively measure a player's experience with computer games.

In a third experiment we developed a method for mathematically modeling emotion using physiological data. A fuzzy logic model transformed four physiological signals (GSR, HR, EMG smiling, EMG frowning) into the emotional dimensions of arousal and valence. A second fuzzy logic model transformed arousal and valence into five emotional states: boredom, challenge, excitement, frustration, and fun. The modeled emotions' means were evaluated with test data, and exhibited the same trends as the reported emotions for fun, boredom, and excitement, but modeled emotions revealed statistically significant differences between three play conditions, while differences between reported emotions were not significant. The details of the fuzzy logic model can be found in [4] while the validation of the model and its potential use in interactive systems can be found in [2].

Implications of results

In the course of our experiments, we have made significant contributions to affective computing and HCI evaluation methodologies, and have extended the applicability of fuzzy logic to a new domain. Specifically, we provided: a systematic exploration of how the body responds to interactive play environments; rules and guidelines for conducting research in this domain; support that physiological measures can be used to objectively measure a player's experience with entertainment technology; and evidence that normalized physiological measures of experience with entertainment technology correspond to subjective reports.

In addition, our models of emotion provide a continuous metric for evaluation. Modeled emotions provide an objective and quantitative approach to evaluating play technologies, but can also be viewed over an entire experience, revealing the variance within a condition, not just the variance between conditions. This is especially important for evaluating user experience with entertainment technology, because the success during play is determined by the process of playing, not the outcome of playing [6]. Continuously representing emotion is a powerful evaluative tool (with a high evaluative bandwidth) that can be easily combined with other methods. Given a time series of emotional output, researchers can use interesting features in the modeled emotion output to index other evaluative data sources such as video or screen captures of the play environment, or other objective behavioural data.

Future Directions

Our models and results are still preliminary and require further investigation before they are robust enough for deployment outside of a laboratory. Most importantly, we need to employ a better method of scaling the physiological responses to deal with the high variability between individuals. In addition, our models were developed *post hoc*; more research may make these models run in real-time with new users unknown to the system.

We are presently considering other emotional states relevant to interaction with entertainment technology (such as disappointment or pride), which can be described by arousal and valence. Other more complex emotions, such as schadenfreude, are difficult to describe in the emotional dimensions of arousal and valence and require more research to model. In addition, we are looking into other physiological sensors such as EEG and pupil diameter.

Finally, we would like to integrate our approach with other behavioural evaluation methods. Details on these future refinements to the emotional models can be found in [2, 4].

References

1. Mandryk, R.M, Inkpen, K.M. (2004). Physiological indicators for the evaluation of co-located collaborative play. *Proceedings of the Conference on Computer Supported Cooperative Work (CSCW 2004), Chicago, USA,* 102-111.
2. Mandryk, R.M., Atkins, M.S., Inkpen, K.M. (2006). A continuous and objective evaluation of emotional experience with interactive play environments. *Proceedings of the Conference on Human Factors in Computing Systems (CHI 2006). Montreal, Canada*, 1027-1036.
3. Mandryk, R.M, Inkpen, K.M., Calvert, T.W. (2006). Using psychophysiological techniques to measure user experience with entertainment technologies. *Behaviour and Information Technology* **25(2)**, 141-158.
4. Mandryk, R.M., Atkins, M.S. (2007). A fuzzy physiological approach for continuously modeling emotion during interaction with play environments. *International Journal of Human-Computer Studies*, **6(4)**, 329-347.
5. Norman, D. A. (2002). Emotion and design: Attractive things work better. *Interactions* **9**, 36-42.
6. Pagulayan, R. J., Keeker, K., Wixon, D., Romero, R., & Fuller, T. (2002). User-centered design in games. In J. Jacko & A. Sears (Eds.), *Handbook for human-computer interaction in interactive systems,* Mahwah, NJ: Lawrence Erlbaum Associates, Inc., 883-906.

Think Aloud-method during fMRI to determine neuronal correlates of subjective experience of video game playing

Martin Klasen[1], René Weber[1,3], Krystyna A. Mathiak[1], Mikhail Zvyagintsev[1], and Klaus Mathiak[1,2]

[1] *RWTH Aachen University Hospital, Aachen, Germany, mklasen@ukaachen.de*

[2] *King`s College, University of London, London, UK*

[3] *University of California, Santa Barbara, CA, USA*

The subjective experience of video games and its consequences have been subject of public debate as well as of scientific research. Recently, neuroimaging has been applied to study the neural correlates of behaviour during video games [1]. Nevertheless, systematic research on game experience in general and the search for its neural correlates in special is still at its starting point and remains a fascinating topic for future investigations. To investigate central dimensions of video game experience we applied Think Aloud (TA), i.e. an introspective method to assess structured retrospective reflections on thoughts and feelings while re-watching one's own playing [2]. Originating from psychological research, the TA was developed from the much older introspection method. Introspection is based on the assumption that a person is able to observe his or her thoughts, feelings, and cognitive processes, as well as one is able to observe processes in the outside world. These skills can also be trained so that the subject can observe them in a systematic way and make them accessible to others by verbalizing them.

In 18 subjects we recorded functional magnetic resonance (fMRI) data during playing a first-person shooter game (Counter Strike: Source). In simple terms, fMRI is a technology in which the amount of blood flow in circumscribed brain areas is calculated on the basis of magnetic responses of hemoglobin. Blood flow in brain areas is closely correlated with neural activity in those areas, and increase or decrease of blood flow therefore indicates increasing or decreasing brain activation, respectively. After some training time outside the scanner room the subjects played a total of three game sessions (12 minutes each) during fMRI measurement. The TA sessions on the recently played game content took place directly after each session inside the scanner. This technique has three main advantages: First, it does not interrupt or disturb the game play itself as it would be the case if the players performed TA simultaneously during the play; second, the player still has the immediate impression of the recently finished session and can perform the TA in the same setting as the game play which facilitates memory effects; and third, fMRI can be measured during the TA session itself, giving insight into neural correlates of speech production, memory, and empathy. Participants were instructed to report whatever they want, but not to fall into a "description mode", i.e., to describe what is visible or audible in the game recording. We instructed them to focus mainly on two aspects: the difficulty of the game (*How hard/easy was the game play in the respective situation?*) and the aspect of subjective game enjoyment (*How much did you enjoy the game play in the respective situation?)*. At the end of each gaming block, the subjects watched a recording of their recently played sequence. They were instructed to verbally report continuously their estimation of difficulty and the amount of enjoyment they had during playing the respective scenes. The speech was recorded with an MR compatible optical microphone built as a prototype from the company Sennheiser Electronics (Wedemark, Germany). The microphone was attached to the radiofrequency head coil and its signal was filtered online to reduce the EPI noise of the scanning sequences. To this end the experimenter could listen to the speech in on-line mode and deliver visual prompts if necessary ("Please speak up!", "Please keep on talking!", and "Please focus on game enjoyment and game difficulty!"). The experimenter registered all commented events but also non-commented game events if they appeared relevant for the challenge and enjoyment of the game.

Since expressing one's thoughts verbally is not a common behavior, a 12 minutes test block outside the scanner took place, including detailed instructions and a training phase. The test block consisted of a 6 minutes playing session followed by a 6 minutes Think Aloud phase on the recently played sequence which could be repeated if necessary. The experimenter ensured that the subject had understood the instructions and was able to complete the task.

Evaluation of the Think Aloud was done by three independent raters (three undergraduate students at the RWTH Aachen University Hospital) and one supervisor independently to avoid subjective interpretation biases. For testing inter-coder reliability we used Krippendorff`s Alpha coefficient [4]. Intensive trainings yielded an overall inter-coder reliability of .70 - .80. The Flow concept [3] guided the content analysis of the game play recordings and of the Think Aloud comments. With statistical parametric mapping (SPM5), we determined neuronal networks involved in major experiential dimensions. Predominant cerebro-thalamic motor and visual networks reflected the categories, but experiential factures such as failure were associated with inhibition of the reward system (e.g., caudate nucleus). Positive and negative appraisal as reported in the TA also revealed circumscribed brain activation patterns. Thus the findings present a neurophysiological observer-independent validation of the method and the underlying psychological methods. We conclude that the Think Aloud method may help to disentangle the link between basic neuronal subsystems and human subjective experience as reflected by Think Aloud by offering a way to assess thoughts and feelings of a person during complex and interactive tasks and providing the basis for a structured, content-oriented data analysis.

References

1. Mathiak, K. & Weber, R. (2006). Toward brain correlates of natural behavior: fMRI during violent video games. *Human Brain Mapping*, 2006 Dec; **27(12)**, 948-56.
2. Van Someren, M. W., Barnard, Y. F., Sandberg, J. A. C. et al. (1994). *The think aloud method. A practical guide to modelling cognitive processes.* London: Academic Press.
3. Csíkszentmihályi, M. (1975). *Beyond Boredom and Anxiety.* San Francisco: Jossey-Bass.
4. Krippendorff, K. (2004). *Content Analysis, An Introduction to Its Methodology*. Thousand Oaks, CA: Sage Publications.

Multi-modal Behavioral Cues from Bodily Interaction in Ambient Entertainment Applications

Anton Nijholt, Betsy van Dijk, and Dirk Heylen

University of Twente, Human Media Interaction (HMI), PO Box 217, 7500 AE Enschede, the Netherlands, a.nijholt@ewi.utwente.nl

Abstract

Exertion interfaces require bodily activity. Users have to perform exercises, they have to dance, they have to golf or football, and they have to train particular bodily skills. Unlike game interfaces where we can observe lots of research activity recently trying to define, interpret and evaluate issues such as 'flow' and 'engagement', in exertion interfaces these concepts need to be reconsidered and new ways of evaluation have to be defined. Here we embed exertion interface research in ambient intelligence and entertainment computing research. Examples are discussed and views on evaluation are expressed.

Introduction

In previous years exertion interfaces have been introduced [1]. In game or entertainment environments the 'user' may take part in events that require bodily interaction with sensor-equipped environments. E.g., in an urban game, mobile devices may be used to inform the users about activities they have to perform or about activities of their partners or opponents in the game. The game can require the gamer to walk, run, or perform other activities, in order to compete or cooperate with others involved in the game. Exertion interfaces have also been introduced in home and office environments that offer, elicit and stimulate bodily activity for recreational and health purposes.

For example, in a smart, sensor-equipped, home environment bodily activity can be employed to control devices, or the smart home environment might anticipate our activities and behave in a pro-active and anticipatory supporting way. Although in home environments there is freedom when and how to perform tasks, there nevertheless are regular patterns of bodily activity and therefore activities can be predicted and anomalies can be detected. In task-oriented environments, e.g. an office environment, people probably have more well-defined tasks where efficiency plays an important role. Smart office furniture can provide context and task aware support to a moving office worker.

Exertion and Entertainment Interfaces

In order to design and implement successful exertion interfaces we need an environment that can detect, measure, and interpret physical activity. One of the best known exertion interfaces is 'sports over a distance', where players from different sites have to hit a wall with a ball [1]. The position on the wall and the force with which the ball hits the wall are important for winning or losing the game. In this particular exertion interface there is no direct sensing of body movements or physiological information. Only the result of the exertion is measured and mediated. There also exist exertion interfaces with direct sensing of bodily activity (body movements, gestures, bodily and facial expressions, dynamic aspects of expression, etc.) and of speech activity that accompanies bodily activity (effort and pain utterances, laughs, prosodic aspects of speech utterances ...). Cameras and microphones allow visual and audio processing of a user's activity and any other sensors are available. There are sensors that provide information about location changes (tracking bodies and faces of individuals), frequency and expressiveness of movements, effort measuring, etc. One step further is to take into account physiological information obtained from the user. This information can be used both to guide the interaction and to measure the user experience. In particular brain-computer interfacing (BCI) is a source of information in our research to learn about the way the user experiences the interaction besides offering the user control.

Multimodal, Joint, and Coordinated Activity in Embodied Interaction

Exertion interfaces emphasize the conscious use of bodily activity (jogging, dancing, playing music, sports, physical exercises, fitness, etc.) in coordination and sometimes in competition with other human users (friends, community or team members, accidental passers-by, opponents, etc.). Real-time coordinated interaction between human partners or between humans and virtual or robotic partners makes exertion interfaces exciting. Coordination may be required by the rules of the game or the exercise or the tasks ask for it, but most of all people engage in coordinated interaction because it brings satisfaction and enjoyment. For users of exertion interfaces the interaction supporting feedback and the interaction experience are important [2].

We take inspiration from Clark's work on joint activity in many of our studies. [3]: "A joint action is one that is carried out by an ensemble of people acting in coordination with each other. As some simple examples, think of two people waltzing, paddling a canoe, playing a piano duet, or making love."

Communication, like dancing, includes coordinated nonverbal activity. We have studied face to face conversations, multi-party interaction, interactions between a virtual and a human dancer [4], a virtual conductor and a human orchestra [5], and a physiotherapist and her student [6]. Underlying joint activities are rules and scripts. To learn these and to put them into practice requires social intelligence, guided by empathy, moods and emotions. Despite many research results from social and behavioral sciences, computational models of joint activities are hardly available. This makes it difficult to design interfaces that aim at providing a similar interactional experience between real humans and virtual humans or robots, as is provided in a real-life human-human exertion activity, as in dancing, paddling, playing quatremains, and making love. Endowing the computer with a human-like appearance strengthens the expectation the computer will take part in joint activities in human-like ways. Hence, there is a need for computational modeling of human joint activities. We replace one of the human partners in a joint exertion activity by a computer (i.e., a robot or a virtual human). Hence, we need to model the exertion interaction in order to have the computer behave in a natural and engaging way.

In addition to rules that underlie joint activity there can be a need to align the interaction to external events over which the interaction partners do not necessarily have control. E.g., if we have a human and a virtual dancer then their moves have to be aligned with the music. Similarly, a virtual conductor and his human orchestra follow the score; a virtual aerobics trainer interaction partners do not necessarily have control. E.g., if we have a human and a virtual dancer then their moves have to be aligned with the music. Similarly, a virtual conductor and his human orchestra follow the score; a virtual aerobics trainer and human student have to align their movements to the supporting music.

In our research we look at

Measuring activity to improve the interaction (using off-line information: history, personality, …) and adapt the system to the user's history and current activities

Measuring activity in order to know about the involvement of the user (flow, pleasure, pain, effort, tiredness, ...) and adapt the system in order to improve engagement

Earlier [7] we argued that for applications that have the interaction itself as their goal, the interaction and the user experience need to be evaluated, rather than efficiency. In our present research we investigate ways to measure engagement by looking at the degree of coordination between the activities of a human and a virtual partner in exertion and other entertainment interfaces [8]. In this research, supported by [9,10] we investigate how to make entertainment interactions more enjoyable by looking at interaction synchrony, where, on the one hand we aim at disturbing this synchrony in order to introduce a new challenge, and on the other hand we aim at convergence towards coordinated anticipatory multi-modal interaction between human and artificial partner and environment.

Acknowledgements. This work has been supported by the GATE project, funded by Dutch NWO and the Netherlands ICT Research and Innovation Authority (ICT Regie).

Refererences

1. Mueller, F., et al. (2003). Exertion Interfaces: Sports over a Distance for Social Bonding and Fun. In: *Proc. CHI 2003*. ACM Press, USA, 561-568.
2. Bianchi-Berthouze, N., et al. (2007). Does body movement engage you more in digital game play? And Why? In: *Affective Computing and Intelligent Interaction.* LNCS 4738, Springer, 102-113.
3. Clark, H (1996). Using Language. Cambridge University Press.
4. Reidsma, D., et al. (2006). Towards Bi-directional Dancing Interaction. *5th Intern. Conf. on Entertainment Computing (ICEC 2006)*, R. Harper et al. LNCS 4161, Springer, Berlin, 1–12.
5. Maat, M. ter, et al. (2008). Beyond the Beat: Modeling Intentions in a Virtual Conductor. *2nd Intern. Conf. on INtelligent TEchnologies for interactive enterTAINment (INTETAIN).*
6. Ruttkay, Z.M. & Welbergen, H. van (2006). On the timing of gestures of a Virtual Physiotherapist, *3rd Central European MM & VR Conf.* C.S. Lanyi (Eds), Pannonian Univ. Press, Hungary, 219-224.
7. Poppe, R.W. et al. (2007). Evaluating the Future of HCI: Challenges for the Evaluation of Emerging Applications. In: *AI for Human Computing,* T. Huang et al. (eds.), Springer, 234-250.
8. Nijholt, A. et al. (2008). Mutually Coordinated Anticipatory Multimodal Interaction. In: Nonverbal Features of Human-Human and Human-Machine Interaction. LNCS, Springer, to appear.
9. F. Tanaka et al. (2004). Dance Interaction with QRIO. International Workshop on Robots and Human Interactive Communication, IEEE.
10. Tomida, T. et al. (2007) miXer: The Communication Entertainment Content by using "Entrainment Phenomenon" and "Bio-feedback". *ACE 2007*, ACM, 286-287.

Psychophysiology of digital game playing: The relationship of self-reported emotions with phasic physiological responses

N. Ravaja and J.M. Kivikangas

Center for Knowledge and Innovation Research, Helsinki School of Economics, Helsinki, Finland, ravaja@hse.fi

Digital game playing is the fastest growing form of entertainment media [1]. Given that emotions and emotion-related variables (e.g., competitiveness) play a critical role in gaming behavior [2, 3], increasing our understanding of what actually happens during video game playing in terms of emotional responses is of increasing importance from the point of view of communication researchers, media psychologists, and game industry (game developers).

Emotions are biologically based action dispositions that comprise three components: subjective feeling, expressive behavior, and physiological component [4]. A dimensional theory of emotion holds that all emotions can be located in a two-dimensional space, as coordinates of valence (ranging from negative to positive) and arousal (or bodily activation; ranging from low to high).

Facial electromyography (EMG) is the primary psychophysiological index of hedonic valence [5]. That is, EMG activity increases with the contractions of the facial muscle groups responsible for positive and negative emotional expressions. It is well established that increased activity at the zygomaticus major (cheek) and corrugator supercilii (brow) muscle regions is associated with positive emotions and negative emotions, respectively, when viewing media [5, 6]. In addition, increased activity at the orbicularis oculi (periocular) muscle area has been associated with positive and high-arousal emotions [5, 7]. Electrodermal activity (EDA), commonly known as skin conductance, is an important psychophysiological index of arousal [5]. As people experience arousal their SNS is activated, resulting in increased sweat gland activity and skin conductance. Several studies using picture stimuli, for example, have shown that EDA is highly correlated with self-reported emotional arousal.

However, the interpretation of psychophysiological measures is highly dependent on the context and research paradigm [5] and they have not been properly validated in the context of digital games. Although there is some evidence that tonic (i.e., averaged across the whole play session) facial EMG and EDA are associated with emotional processes during digital game playing [3], it is not clear whether they can be used to index emotional responses to in-game events. Therefore, in this paper, we examine the relationship of phasic facial EMG and EDA responses to different in-game events with self-reported emotions elicited by these events.

Methods

Participants were 40 (21 male and 19 female) volunteering Finnish young adults, who ranged from 18 to 31 (mean = 22.7) years of age and played digital games at least four hours per month.

The game used was FUGAmod, a custom-made mod based on popular first-person shooter (FPS) game Half-Life 2 (Valve Corporation, Bellevue, WA, 2004), played on a powerful desktop computer. According to the story conveyed to the players beforehand, the player character (PC) is a freedom fighter in a dystopian future that needs to rescue some people (allies), one by one, from an occupied building and distinguish these from enemy spies. The facial expression of the non-player characters (NPCs) was also varied. Accordingly, there were (a) smiling allies that should be rescued, (b) smiling spies that should be killed, (c) frowning allies that should be rescued, and (d) frowning spies that should be killed. Each of these NPC subtypes occurred three times during the task, so there were a total of 12 NPC encounters in the entire play period. The game automatically sent codes for different in-game events to the psychophysiological data acquisition system via a serial cable. The played game was stored as digital video. After playing the game, the participants saw 6-s video clips of the in-game events and self-rated their emotional responses on the valence and arousal dimensions to each event.

During the game, skin conductance was recorded with the Psylab Model SC5 24 bit digital skin conductance amplifier that applied a constant 0.5 V across Ag/AgCl electrodes with a contact area of 8 mm diameter (Med Assoc. Inc., St. Albans, VT). Facial EMG activity was recorded from the left corrugator supercilii, zygomaticus major, and orbicularis oculi muscle regions as recommended by Fridlund and Cacioppo [8], using surface Ag/AgCl electrodes with a contact area of 4 mm diameter (Med Assoc. Inc., St. Albans, VT). The raw EMG signal was amplified, and frequencies below 30 Hz and above 400 Hz were filtered out, using the Psylab Model EEG8 amplifier. The digital data collection was controlled by Psylab7 software, and the signal was sampled at a rate of 500 Hz.

Mean values for the psychophysiological measures were derived for one 1-s epoch before each event (this provides a local baseline; Second 1) and for two 1-s epochs after event onset (Seconds 2 and 3). Contrast scores for the linear and quadratic trends across Seconds 1-3 in EMG activity were calculated. The data were analyzed by the Linear Mixed Models procedure in SPSS with restricted maximum likelihood estimation and a first-order autoregressive covariance structure for the residuals.

Results

Table 1 shows the results of the analyses of the relationship of phasic psychophysiological responses to different in-game events with self-reported valence and arousal. It was found that the linear trend across Seconds 1-3 in zygomatic EMG activity (contrast score) was positively associated with self-reported pleasure, $p = .047$. Likewise, the linear trend across Seconds 1-3 in orbicularis oculi EMG activity was positively related to self-reported pleasure, $p = .010$. In addition, the linear trend across Seconds 1-3 in corrugator supercilii EMG activity was negatively related to self-reported pleasure, $p = .003$. Phasic EDA responses were not related to self-reported arousal.

Conclusion

The present study showed that phasic zygomaticus major and orbicularis oculi EMG responses to different events in a FPS game were positively associated with self-reported pleasure elicited by these events when re-watching them. In addition, phasic corrugator supercilii EMG responses to the in-game events were negatively associated with self-reported pleasure. This study provides evidence for the validity of phasic facial

Table 1. *Results of Linear Mixed Models Analyses: Relationship of Self-reported Valence and Arousal with Phasic Physiological Responses to Game Events*

Variable Source	df	F	p
ZM EMG[a]			
Intercept	1, 181.06	2.480	.117
Event type	7, 720.71	.837	.556
Event number[b]	2, 721.57	1.983	.138
Valence[c]	1, 303.77	3.985	.047
Arousal[d]	1, 58.58	.077	.782
CS EMG[a]			
Intercept	1, 328.43	4.336	.038
Event type	7, 747.56	.648	.716
Event number[b]	2, 748.52	.647	.524
Valence[c]	1, 644.82	8.966	.003
Arousal[d]	1, 139.02	.005	.945
OO EMG[a]			
Intercept	1, 231.57	3.887	.050
Event type	7, 719.93	1.225	.286
Event number[b]	2, 725.48	.516	.597
Valence[c]	1, 396.83	6.705	.010
Arousal[d]	1, 82.74	.165	.686

Note. EMG = electromyography; CS = corrugator supercilii; OO = orbicularis oculi; ZM = zygomaticus major.
[a]*Linear trend across Seconds 1 through 3 (contrast score)*
[b]*Sequence number of an event*
[c]*Self-reported valence*
[d]*Self-reported arousal*

EMG responses to in-game events as indices of emotional valence. In addition to being of interest to game researchers and media psychologists, we suggest that information on phasic emotional reactions to game events and event patterns can be applied in game design. Phasic EDA responses were, however, not related to self-reported arousal. This questions the validity of phasic EDA responses as measures of arousal during digital game play.

References

1. Holson, L.M. (2004, April 10). Out of Hollywood, rising fascination with video games. *The New York Times on the Web*. Retrieved April 10, 2004, from http://www.nytimes.com/2004/04/10/technology/10GAME.html
2. Grodal, T. (2000). Video games and the pleasures of control. In D. Zillmann & P. Vorderer (Eds.), *Media entertainment: The psychology of its appeal* (pp. 197--212). Mahwah, NJ: Lawrence Erlbaum Associates.
3. Ravaja, N., Saari, T., Salminen, M., Laarni, J., & Kallinen, K. (2006). Phasic emotional reactions to video game events: A psychophysiological investigation. *Media Psychology,* **8**, 343-367.
4. Lang, P.J. (1995). The emotion probe. Studies of motivation and attention. *American Psychologist,* **50***,* 372-385.
5. Ravaja, N. (2004a). Contributions of psychophysiology to media research: Review and recommendations. *Media Psychology,* **6***,* 193-235.
6. Ravaja, N., Kallinen, K., Saari, T., & Keltikangas-Järvinen, L. (2004). Suboptimal exposure to facial expressions when viewing video messages from a small screen: Effects on emotion, attention, and memory. *Journal of Experimental Psychology: Applied,* **10***,* 120-131.
7. Ravaja, N., Saari, T., Kallinen, K., & Laarni, J. (2006). The role of mood in the processing of media messages from a small screen: Effects on subjective and physiological responses. *Media Psychology,* **8**, 239-265.
8. Fridlund, A. J., & Cacioppo, J. T. (1986). Guidelines for human electromyographic research. *Psychophysiology,* **23***,* 567-589.

Towards real-time behavioral indicators of player experiences: Pressure patterns and postural responses

Wouter van den Hoogen, Wijnand IJsselsteijn, Yvonne de Kort, and Karolien Poels
Game Experience Lab, Human Technology Interaction Group, Eindhoven University of Technology, Eindhoven, The Netherlands, W.M.v.d.Hoogen@tm.tue.nl

Introduction

The current paper sets out to describe a first exploration of behavioral expressions that could serve as real-time indicators of experiences related to playing digital games. In this paper, we focus primarily on pressure patterns exerted on a physical control device, and postural responses. Based on this exploration, we present our progress in developing a set of behavior-based measures of such player experiences and their validation through an experimental study.

Behavioral indicators of flow, frustration and boredom

Enjoyment is arguably the single most important motivation for people to play digital games. It is a factor that game developers would like to understand more fully in order to optimize game design [6]. In his pioneering work, Csikszentmihalyi [1,2] found that many elements of enjoyment are universal. During optimal experiences, people are completely absorbed by an activity, losing track of time and a preoccupation with self. Through striking the balance between a person's skills and the challenges an activity offers, that person may arrive in a psychological state known as *flow*. Although Csikszentmihalyi's work on flow suggests it is quite a rare experience, it certainly is a state that rings familiar to players of digital games. Flow may gradually increase over the course of the game in a homeostatic positive feedback loop, until either the challenge becomes too great (resulting in frustration) or the player's skill outpaces the challenges the game can offer (leading to boredom). Being able to measure in real-time when a game is becoming unacceptably boring or frustrating is likely to contribute to a more optimal player experience, as such information can be provided directly to the game engine, allowing it to adjust the game dynamically to the player's psychological state.

Behavioral responses exhibited during gameplay, such as postural movements, gestures, facial expressions, or pressure exerted on interface devices, constitute a potentially promising class of measures in this regard. Mota and Picard [4] have, for instance, demonstrated that postural patterns can be indicative of learner interest. They developed a system to recognize postural patterns and associated affective states in real time, in an unobtrusive way, from a set of pressure sensors on a chair showing that the dynamics of postures can distinguish with significant reliability between affective states of high interest, low interest and boredom, each relevant to a gaming situation as well.

Research by Mentis and Gay [3] and Park, Zhu, McLaughlin and Jin [5] provides evidence that the force people apply to interface devices can be interpreted as an indicator of negative arousal. Indeed, Park et al. [5] showed a correlation between higher pressure patterns and facial expressions showing negative affect, thereby providing converging evidence that pressure exerted may be related to frustration. In the domain of digital games, Sykes and Brown [7] found that the mean pressure exerted by players on a gamepad's button increases with the difficulty level of a game. However, their results are ambiguous in that higher pressure may be associated with either a positive emotion (more excitement) or a negative one (more frustration), both of which can plausibly occur at higher difficulty levels. In sum, the literature provides tentative evidence that behavioral responses to game events may be fruitfully explored as real-time measures of player experience, in particular interest, boredom and frustration.

Construct validity study

In order to test the utility and validity of behavioral indicators of player experiences, we carried out an experiment in which thirty-two participants played 3 customized levels of a first-person shooter (Half Life 2). The levels varied in difficulty with the aim to induce boredom (easy level), enjoyment (moderate level), and frustration (hard level). The experiment had a within-subjects design, counterbalancing the order in which the levels were played. Dependent measures included several real time behavior measurement systems, including a pressure-sensitive chair, inspired on the work of Mota and Picard [6], and a pressure sensitive mouse and keyboard (see Figure 1). The chair was designed such that changes in forward-backward and sideways movement can be sensed. Complementing these systems we also employed a 3-axis Phidgets accelerometer measuring tilt and acceleration of participants' upper body in both the frontal and lateral plane. In addition, observational coding of sitting position was done, based on video recordings of each participant. Moreover, several self-report measures were applied (GEQ; SAM; FlowGrid) in order to ensure that the behavioral data could be sensibly interpreted, and to check convergent validity of the behavioral indicators with such self-report data.

The results show a consistent pattern between self report, observed behavioral indicators and automatically captured measurements. The self report measures indicate frustration for the hard level, balanced play for the moderate level, and less engagement at the easy level. In line with literature and our hypotheses, behavioral measures for movement (e.g. accelerometer data and observational coding of movement) and force (e.g. on the mouse) appeared to be indicative of frustration in the hard level. More specifically, mean accelerometer data was highest in the hard level indicating stronger movement. Similarly, data from both the automatic chair and observational coding show the range in forward-backward movement to be significantly higher in the hard level as compared to the other two levels. The force measurements of the mouse show that participants applied more force on the mouse buttons (both maximum and mean over the level) in the hard level, an effect consistent with previous findings [7]. Our findings further show maximum force on the mouse and range in sitting position to be correlated with the self-report measurement of frustration, providing additional, convergent validity for the measurement of frustration through these behaviors.

To conclude, the first analysis of our results bode well for the applicability of behavioral indicators to measuring player experiences in relation to digital games. However, our pressure and posture measures appear to be particularly sensitive to frustration but not so much to boredom. This may be due to our manipulation strength, the sensitivity of our set

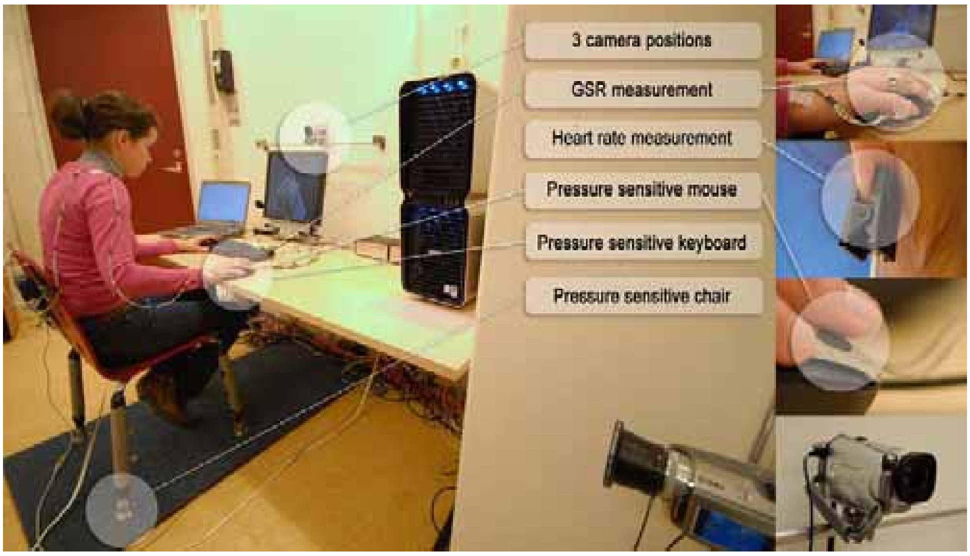


Figure 1. *Measurement systems as used in the experiment.*

of measures, the relation between boredom and behavior, or a combination of these factors. This issue will be addressed through further analysis of our current findings, and future experiments.

References

1. Csikszentmihalyi, M. (1975). *Beyond Boredom and Anxiety*. San Fransisco: Jossey-Bass.
2. Csikszentmihalyi, M. (1990). *Flow. The Psychology of Optimal Experience*. New York: Harper & Row.
3. Mentis, H.M. and Gay, G.K. (2002). Using touchpad pressure to detect negative affect. *Proceedings of Fourth IEEE International Conference on Multimodal Interfaces* 2002, 406 - 410
4. Mota, S. and Picard, R.W. (2003). Automated Posture Analysis for Detecting Learner's Interest Level. Workshop on Computer Vision and Pattern Recognition for Human-Computer Interaction (CVPR HCI). Available: http://affect.media.mit.edu/pdfs/03.mota-picard.pdf
5. Park, N., Zhu, W., Jung, Y., McLaughlin, M., and Jin, S., (2005). Utility of haptic data in recognition of user state. *Proceedings of HCI International 11*. Lawrence Erlbaum Associates. Available: http://imsc.usc.edu/haptics/paper/manuscript_hcii2005_final.pdf
6. Phillips, B. (2006). Talking about games experiences – A view from the trenches. *Interactions, **13*** (5), 22-23.
7. Sykes, J. and Brown, S. (2003). Affective gaming. Measuring emotion through the gamepad. *ACM CHI* 2003, 732-733.

Affect Decoding Measures and Human-Computer Interaction

L.D. Riek, S. Afzal, and P. Robinson
Computer Laboratory, University of Cambridge, Cambridge, UK, lr314@cl.cam.ac.uk

Introduction

Affect measurement has been identified as a critical issue for Human-Computer Interaction (HCI) [4]. There is an established body of work in the communication science and psychology literature considering an individual's ability to decode affect. However, this area has not yet been applied to HCI. Decoding affect is defined as "the ability to sense, perceive accurately, and respond appropriately to one's personal, interpersonal, and social environment" [2]. Objective measurement of affect decoding is relevant for a variety of HCI applications, such as designing and selecting distance collaboration tools, analysis of behavioral video, and explicitly describing observed nonverbal behavior during dyad and group interaction studies. We are used to adapting interfaces for users with disabilities [6]; it is equally important that systems that support cooperative work should adapt to their users' difficulties with decoding affect. It is therefore necessary to be able to measure an individual's affective decoding abilities.

Measurement Methods for Decoding Affect[1]

A wide range of methods are available to quantitatively measure how people decode affect. These methods usually involve evaluation of one or more channels of affective communication, such as body gesture, facial expression, vocal behavior, or other nonverbal cues. Each method has its advantages and drawbacks; indeed, when one reviews such methods it is useful to note the internal consistency score of the method. We describe four methods that measure the affect decoding ability of subjects presented with non-static stimuli. The stimuli presented are typically video and/or audio of people expressing various levels of affect.

Profile of Nonverbal Sensitivity Test (PONS) [11]

PONS measures the accuracy with which subjects infer the nonverbal cues of emotionally laden content acted out by a female expresser via face, body, and vocal channels. The test is available in both full and short versions. The full-length test consists of 220 2-second audio clips, video clips, or both. After each clip, subjects are given a multiple choice question and asked to select the correct portrayal. People with higher PONS scores are rated as more interpersonally sensitive by supervisors and peers, have more well-adjusted personalities, and are rated as better job performers as clinicians and teachers [5].

Communication of Affect Receiving Ability (CARAT) [3]

CARAT is a standardized method for measuring how subjects decode spontaneously generated facial expressions. First, stimulus subjects are videotaped while viewing 32 emotionally evocative slides (i.e., laughing children or injured animals). Test subjects then view these recorded segments and are asked to choose which slide the stimulus subject was viewing. The total score is a measure of the test subjects' emotional decoding skill.

Empathic accuracy standard-cue methodology [7]

This test measures the "empathic accuracy" of subjects as they watch videos of three female patients interacting (individually) with a therapist. Subjects are shown and/or played 30 15s video clips. After each clip, they are asked to rate if the patient was feeling something and, if so, to describe the feeling in a single sentence. Following this, eight independent raters then compare the subjects' response in tandem with the video and measure accuracy on a three-point scale. An emphatic accuracy score is then generated [7].

Empathy Quotient (EQ) [1]

EQ is another measure of empathy and is a short, easy to use and score evaluative method. Subjects are given an 80-question, 4-point Likert scale pencil-and-paper test. EQ has 40 questions that probe emotional empathy ("It doesn't bother me too much if I am late meeting a friend."), and 20 filler questions. On the empathic behavior questions subjects receive 1 point for a mild response and 2 points for a strong response. The empathy questions are evenly balanced between "strongly agree" and "strongly disagree" responses to avoid bias. EQ has been shown to be significantly lower in adults with Asperger Syndrome or high-functioning autism compared with controls. Further, women tend to score significantly higher on EQ than men [1].

Applications to HCI

Methods for measuring affective decoding ability are applicable to a variety of tasks relevant to the HCI community. First, for Computer-Supported Collaborative Work (CSCW) designers, the ability to design systems based on the affective decoding scores of their target user population could greatly inform their decision-making. Second, when HCI practitioners are selecting subjects for user trials, it can be very important to know their subjects' ability to decode affect. Finally, for scientists who analyze behavioral video data, knowledge of their own affective decoding skills can be helpful in both determining their level of expertise and in explicitly describing how they derive their conclusions.

Sensitivity to behavioral cues is a key factor for efficient coordination and collaboration. For example, a group of non-collocated architects collaborating on a design task involves interpersonal sensitivity. Each architect's ability to understand the affect of their colleagues can greatly impact the success of the task. By pre-testing the architects, a technology designer or selector can better compensate for the strengths and weakness of the group at large. In this instance, the use of the PONS test might prove helpful to provide a multi-channel score with which to work.

Another application of these methods is for video analysis of natural corpora. Affective computing [9] methods often involve dealing with naturally collected data. Labeling this data is important as it serves as the ground-truth for validation of the developed computational techniques. However, training annotators to do this task is often very costly, particularly if they are not skilled at recognizing nonverbal affect. It may be beneficial to use measures such as CARAT, Empathic Accuracy, or EQ to pre-screen annotators, as well as to indicate where training is required. Annotator scores can also serve as reliability indicators for the labeled data.

More generally, for scientists studying affect among dyads or groups, familiarity with measures such as PONS can be very informative for explicitly describing observed nonverbal behavior as well as for communicating how such descriptions were derived.

References

1. Baron-Cohen, S., Wheelwright, S. (2004). The Empathy Quotient: An Investigation of Adults with Asperger Syndrome or High Functioning Autism, and Normal Sex Differences. *Journal of Autism and Developmental Disorders* **34:2**, 163-175.
2. Bernieri, F. J. (2001). Toward a Taxonomy of Interpersonal Sensitivity. In J. A. Hall & F. Bernieri (Eds.). *Interpersonal Sensitivity: Theory and Measurement* pp. 3-20. Mahwah: Earlbaum.
3. Buck, R. (1979). Measuring Individual Differences in the Nonverbal Communication of Affect: The Slide-Viewing Paradigm. *Human Communication Research* **6**, 47-57.
4. Crane, E. A., Shami, N. S., Peter, C. (2007). Let's Get Emotional: Emotion Research in Human Computer Interaction. *In Proc. of Conference on 2007 ACM Conference Computer-Human Interaction (CHI)*.
5. Hall, J. Bernieri, F. J., Carney, D. R. (2005). Nonverbal and Interpersonal Sensitivity. In Jinni, A. Harrigan, J. A., Rosenthal, R., and Scherer, K. R. (Eds.) *The New Handbook of Methods in Nonverbal Behavior Research,* pp. 267-269. Oxford: Oxford University Press.
6. Hughes, G., Robinson, P. (2006). Lecture Adaptation for Students with Visual Disabilities Using High-Resolution. *In Proc. of 2006 ACM Conference on Computers and Accessibility (ASSETS)*.
7. Ickes, W., Marangoni, C., Garcia, S. (1997). Studying Empathic Accuracy in a Clinically Relevant Context. In W. Ickes (Ed.), *Empathic Accuracy*, pp. 282-310. New York: Guilford Press.
8. Jinni, A. Harrigan, J. A., Rosenthal, R., Scherer, K. R. (2005). *The New Handbook of Methods in Nonverbal Behavior Research*. Oxford: Oxford University Press.
9. Picard, R.W. (1997). *Affective Computing*. Cambridge: The MIT Press.
10. Riggio, R. E. (2006). Nonverbal Skills and Abilities. In V. Manusov, V. and Patterson. M. (Eds.), *The SAGE Handbook of Nonverbal Communication*, pp. 79-87. Thousand Oaks: Sage Press.
11. Rosenthal, R.E., Hall, J. A., DiMatteo, M. R, Rogers, P. L., and Archer, D. (1979). *Sensitivity to Nonverbal Communication: The PONS Test.* Baltimore: Johns Hopkins University Press.

[1]For a more thorough overview of these methods, as well as a description of many others, please see *The SAGE Handbook of Nonverbal Communication* [10] and *The New Handbook of Methods in Nonverbal Behavior Research* [8].

Hats off to LiLiPUT:
Experiences with Lightweight Lab Equipment for Portable User Testing

Peter Fröhlich, Peter Reichl, and Antitza Dantcheva
Telecommunications Research Center (ftw.), Vienna, Austria, froehlich@ftw.at

Summary

We share hands-on experiences with LiLiPUT (Lightweight Lab Equipment for Portable User Testing), a custom-designed wearable user testing system. In three Mobile HCI studies, we evaluated the degree to which LiLiPUT currently meets the requirements for mobile applications testing and research. Learnings and suggestions for further improvement are provided.

Towards mobile user testing equipment – requirements and approach

Our notion of computer use is quickly diversifying beyond the classical desktop-based setting. With the transition towards mobile and ubiquitous computing, it is important to also think about new ways of user-centered research. Whereas for web sites and office software, a broad choice of validated design guidance is available, methodologies for design and evaluation of mobile applications and services are still unsatisfactory. On the other hand, for many recent ubiquitous applications and services, the classical stationary usability lab is simply not any more the adequate research environment.

Mobile user testing equipment for telecommunications applications has to fulfill a number of challenging requirements (compare [2]). Most obviously, it needs to be portable. This requirement imposes strong demands on the form factor and power consumption of cameras and recording devices. Mobile user testing must also take account of the various context factors affecting the user cognition and behavior, such as the environment currently visible, weather conditions, the current location, or the noise level. The recordings need to have the same richness, accuracy, and reliability as standard stationary labs. For example, the mobile device screen status should be recordable. Additionally, one of the most important requirements is to enable a natural and seamless usage behavior, without impeding the user by too much observation technology. For instance, the mobile device should be freely movable in every direction and the observation equipment should not be experienced as heavy. It is quite evident that today's solutions for user testing only partly fulfil these requirements (compare [4]).

ftw.'s wearable mobile user testing prototype LiLiPUT (ibid) is designed with the ambitious goal to satisfy all of the above requirements. As Fig. 1 depicts, all the user wears is a hat, which is equipped with small video cameras and a microphone. The captured data is transferred via wireless link to the accompanying observant's backpack. All data – four videos showing the front view, the face view, the remotely captured mobile screen, and the observant's camera, as well as the microphone input from the user's hat and the observer's shoulder – is mixed into one video file and saved on a standard dual core high-speed notebook. In this way, LiLiPUT is able capture much of the mobile context in an accurate and efficient way, while still facilitating a seamless and natural usage situation.

Experiences with LiLiPUT: three case studies

We have used the LiLiPUT prototype throughout the last 18 months in several research and consulting studies to validate our general approach, to investigate its application potential, and to further refine the system design. While of course significantly advancing in terms of technical maturity (e.g. how to avoid noise in wireless transmission in urban environments), we especially wanted to assess users' reactions towards LiLiPUT, as compared to more standard research environments. In the following, three case studies are outlined in order to exemplify our learnings.

A field experiment on contextual factors of Mobile Interactive TV

Contextual factors have a strong influence on how a mobile application is used and perceived. LiLiPUT was used to investigate the user experience of Mobile Social TV, a new application concept which aims at enabling a joint TV viewing experience for remote viewers by using text and audio chat [5]. Three typical mobile situations were of special interest: sitting in a crowded café, walking along a street, and standing at a bus stop. It was evident that – apart from the reactions to Mobile Social TV - these mobile situations also had a different effect on the handling and user perception of LiLiPUT itself. Compared to the stationary café setting, the walking situation required much more attention of the observer to orient herself towards the test subject, in

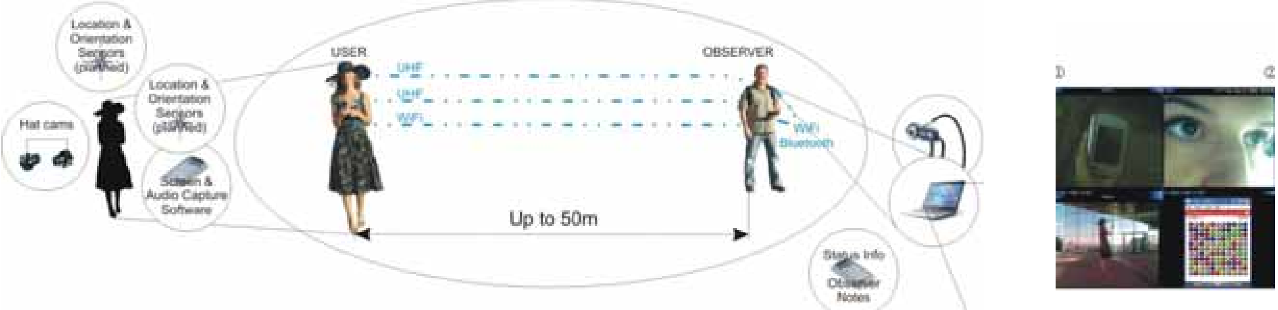


Figure 1. *Architecture of the LiLiPUT System*

order to ensure that all important information is captured. The monitor of a small backup recorder helped us to quickly adjust to the best position. We also learned that weather conditions should strongly taken into consideration in test preparation: the availability of umbrellas and further utensils should be ensured well in advance. When asked about the hat while walking, all test participants felt comfortable. However, in the café, when people are close to each other and social awareness is higher, users mentioned to be embarrassed by wearing the camera-mounted hat.

An outdoor experiment with a ubicomp application

In an outdoor study, we compared different visualization methods for accessing nearby points of interest with a mobile device (see [3] for a description of the visualization methods). While undertaking a 2 hours tour through the city together with the test facilitator, the test participants were asked to complete tasks with different visualization versions. In order to evaluate the subjective experiences towards LiLiPUT, the tests were partly conducted with typical mobile HCI equipment (using a standard DV camera) and the LiLiPUT prototype. When asked in the final interview, it is notable that all participants rated both methods very positively. They neither felt disturbed in using the mobile application by the test equipment (all gave the best rating of 7), nor did they report to be embarrassed by being filmed or wearing the LiLiPUT hat in public streets (mean values of 6.6 and 6.3). Participants did not feel physically annoyed by the hat during the first part of the test (6.75). A drawback, however, is that in the final phase of the 2 hours test, this decreased significantly (3.5), apparently caused by muscular tiring effects. In later test sessions, these problems could be relieved by putting the hat off during the change of test setups, e.g. when walking from one test point to the next one.

A comparative indoor user study on the acceptance and quality of 3G mobile video streaming services

Unexpectedly, although explicitly intended for outdoor studies, LiLiPUT was has also found considerable appreciation for studies in the stationary lab. In a comparative study on the ease of access and the quality of service of circuit-switched and packet-switched mobile video streaming systems [1], LiLiPUT enabled the efficient and simultaneous analysis of the mobile screen display, the test person’s facial expressions, and the overall situation. However, in these stationary testing situations, the weight of the hat seems to have more subjective significance for users than in mobile situations.

Conclusions

We are convinced that wearable user testing equipment is a highly important building block for conducting valid Mobile HCI user studies. The case studies reported above should be regarded as a starting point towards a broad and systematic investigation of future behavioral measurement methods.

References

1. Davies, M., Dantcheva, A., and Fröhlich, P. (2008). Comparing Access Methods and Quality of 3G Mobile Video Streaming Services. To appear in *CHI 2008 Extended Abstracts*
2. Fröhlich, P., Reichl, P., Schatz, R., Baillie, L., Weinberger, W., and Hammer, F. (2006). LiLiPUT - Lightweight Lab Equipment for User Testing in Telecommunications. *HCI 2006*.
3. p2d homepage: http://p2d.ftw.at/applications.html
4. Reichl, P., Fröhlich, P., Baillie, L., Schatz, R., and Dantcheva, A. (2007). The LiLiPUT Prototype: A Wearable Lab Environment for User Tests of Mobile Telecommunication Applications. *CHI2007 Ext. Abstracts*
5. Schatz R, Egger S. (2008). "Social Interaction Features for Mobile TV Services", *Proc. BTS 2008*

Sharing Experiences through Interactive Media: new ways of studying social concepts

A. Visser[1,2] and I. Mulder[2]
[1]User System Interaction, Technical University of Eindhoven, Eindhoven, the Netherlands, a.visser@tue.nl
[2]Telematica Institute, Enschede, the Netherlands, ingrid.mulder@telin.nl

"To have joy one must share it, happiness was born a twin" [Lord Byron]

Sharing of experiences is a way for people to satisfy their need for social interaction. Social interaction is essential for needs of belonging, love and esteem [4]. However, measuring social experiences is hard due to the internal and moment-dependent nature of emotions. This paper outlines different possibilities to measure social experiences when using interactive media. The goal of current research is to design and evaluate a tool that measures certain social concepts related to sharing experiences in a non-intrusive manner.

Measuring social experiences

Questionnaires are often used for measuring social concepts present in social experiences; see for example questionnaires that measure social presence, connectedness and awareness [4]. Questionnaires are usually used for measurement *after* the experience, which can cause cognitive biases [9]. The same problem occurs when using interviews and focus groups. Another method for gathering social user experience data is observation [9]. Although, in comparison to questionnaires, observation is not relying on self-reports from users, it can be highly obtrusive. Also, without any form of self-report from the user conclusions are purely based on behavior which may not be enough to obtain a reliable representation of the user experience. A diary study is a form of indirect observation (and therefore less obtrusive) where users can be asked to write down behavior as well as experiences. Psycho-physiological measurements (e.g., measure pupil size, heart rate, and facial expressions) can measure experiences without asking users any questions, which are difficult to establish in the field [7]. Moreover, a realistic social network for measuring social user experiences is hard to establish in a lab.

Alternatively, the experience sampling method (ESM) has been developed to collect *in-situ* subjective feedback from users [1]. Using this method, users fill out several brief questionnaires during the day by responding to alerts. Because ESM is not relying on recall of experiences, but on current feelings and activities, this reduces cognitive biases. ESM can be quite interruptive in the lives of users, and users need to have the motivation to keep answering the questions [8].

Getting around problems of intrusiveness, self-report, and measuring in the field, three context-aware measurement tools are developed for measuring the user experience. The Context-Aware Experience Sample tool asks users only questions on moments and during activities of interest [5], whereas MyExperience [2] and SocioXensor [6] use mobile devices to collect both objective and subjective data. Although these tools aim at minimizing problems with recall self-report and obtrusiveness, these problems are to a lesser extent still existing. Ideally, researchers would like to measure subjective data *during* an experience without disturbing the experience. Measuring experiences in a social context, makes these measurement issues even more complex because of the influence of social feelings people have in various relationships. Therefore, designing a tool for in-situ measurement of social concepts during moments of sharing experiences in a non-intrusive manner is quite challenging.

Measurement tool

To accept this challenge, requirements have to be considered concerning *which* social concepts to measure and concerning *how* to measure these concepts. Which concepts to measure exactly depends on which social feelings play a role when people are sharing experiences. Hereto, an explorative and inspirational probing diary study (inspired on cultural probes [3]) was conducted to get insight in how users reflect on their sharing experiences: when, how, with whom, and what experiences people share or would like to share, and what emotions people have [10]. Existing literature on measuring experiences (as briefly outlined above) shows that to measure social concepts when sharing experiences, the tool requires to measure **continuously** (to catch the dynamic nature of experiences) which implies **measuring during the experience** (to avoid recall problems), to take **little effort and time**, to be **fun to use** (to keep users motivated), to be able to **measure information about the users social context automatically and unobtrusively**, and to be **mobile** or **wearable** (it can stay and go wherever users are). The current version of the tool will be validated in the upcoming months in a study on sharing experiences while watching social interactive TV. Results from the iterative design and evaluation are to be presented and reflected upon.

References

1. Csikszentmihalyi, M., Larson, R. (1987). Validity and Reliability of the Experience-Sampling Method. *Journal of Nervous and Mental Disease* **175**, 526-536.
2. Froehlich, J., Chen, M., Consolvo, S., Harrison, B., Landay, J. (2007). MyExperience: A System for In Situ Tracing and Capturing of User Feedback on Mobile Phones. *Proceedings of MobiSys 2007*, San Juan, Puerto Rico.
3. Gaver, W., Dunne, T., Pacenti, E. (1999). Cultural Probes. ACM Interactions, **6**(1): 21-29.
4. IJsselsteijn, W.A., van Baren, J., van Lanen, F. (2003). Staying in touch: Social presence and connectedness through synchronous and asynchronous communication media. In: C. Stephanidis, J. Jacko (eds.), *Human-Computer Interaction: Theory and Practice (Part II)*, 924-928.
5. Intille, S. S., Rondoni, J., Kukla, C., Ancona, I., Bao, L. (2003). A context-aware experience sampling tool. *In CHI '03 Extended Abstracts*. ACM Press, New York, 972-973.
6. Mulder, I., Kort, J. (2008). Mixed emotions, mixed methods: the role of emergent technologies to study user experience in context. In: S. Hesse-Biber, P. Leavy (eds.) *Handbook of Emergent Methods in Social Research* (601-612). New York: Guilford Publications.
7. Partala, T. (2005). Affective Information in Human-Computer Interaction. Dissertations in Interactive Technology, Department of computer science, University of Tampere, Finland.
8. Scollon, C.N., C. Kim-Prieto, E. Diener (2003). Experience Sampling: Promises and Pitfalls, strengths and weaknesses. *Journal of Happiness Studies* **4**, 5-34. Kluwer Academic Publishers.
9. Sharp, S., Rogers, Y., Preece, J. (2006). *Interaction design: beyond human-computer interaction.* West Sussex: John Wiley & Sons Ltd.
10. Visser, A., Mulder, I. (accepted). Gaining Insight in (Assessing) Sharing Experiences by means of a Probing Diary Study. Paper for CREATE 2008, London, 24-25 June, 2008.

An Integrated method for a qualitative and quantitative analysis for an ergonomic evaluation of home appliances

M. Romero[1], M. Mazzola[2], F. Costa[3], and G. Andreoni[4]
All the authors belong to Dipartimento INDACO Politecnico di Milano,Via Durando 38/A - Milano, Italy,
[1]maximiliano.romero@polimi.it, [2]marco.mazzola@mail.polimi.it, [3]fiammetta.costa@polimi.it, [4]giuseppe.andreoni@polimi.it

Introduction

The traditional Ergonomics evaluation of the man-product interaction has often been pursued approaching the measurements problem in different ways. Anthropometry, ethnography, video movement analysis, postural analysis and the new quantitative motion capture devices provide several ergonomics variables that have never been integrated. The aim of this paper is to describe a new measurements approach in ergonomics for an integrated evaluation of the interaction between human and product: the method is based upon ethnographic qualitative analysis [1] and a quantitative evaluation of physical biomechanical parameters [2]. The refrigerator is assumed as case study.

Qualitative Analysis

We describe an ethnographic procedure for the qualitative evaluation of the fridge usability based on the users observation in their daily living conditions.

The User-Appliance interaction in a real life context is influenced by the environment (kitchen characteristics, furniture, windows, ecc). The real context make the comparison between observation difficult, but it provides a natural condition of use for the subject [3].

As our approach is “Design for All”[4], we have observed five people with different characteristics:

1. Female, 33 years old, 50 %tile height, without pathologies.
2. Female, 65 years old, 3%tile height, with respiratory disease.
3. Female, 50 years old, 50 %tile height, without pathologies.
4. Female, 50 years old, 50 %tile height, hypo seeing.
5. Male, 40 years old, paraplegic user (in wheelchair).

Observation instruments

We used videocamera recording for the task performance evaluation and voice recording for the subjective interview. One plastic shopping bag was filled with seven products.

The shopping bags contained the same products for each subject. The products were chosen in order to invite the subjects to use all the particular locations in the fridge (Packaging with 6 eggs, 1,5lt bottle of water, a big plant of lettuce, frozen peas, 1lt bottle of milk, butter and yogurt)

Observation set-up

The ethnographic observations were realized meeting the people in their own houses and asking them to execute a task. After this we interviewed the subjects about the fridge’s usability.

Five subjects (selected by age and mobility skills) were observed. All the considered subjects are skilled into the product utilization in their daily activities.

The task was divided in 3 groups of activities:

1. Refilling products
2. Washing the vegetable's drawer
3. Setting the temperature.

Each group of activities is composed of singles free activities (Pull down the shopping bag, Open the fridge door, positioning the products, etc). We left the people use their own natural strategy for the task performing.

Each performance lasted about 15 minutes.

Video evaluation

Five videos were evaluated in order to compare different strategies and times of performance. They were divided in groups and activities, as previously said. Special attention was paid to biomechanical strategies and it was analyzed in order to define the different strategies of movement.

Interview evaluation

The interview was driven by 10 open questions (What are best / worst things of your fridge? Do you have some problem during wash the drawer? How many times you change the temperature?, etc) The answers helped us to define the users profile and their relationship with the fridge.

Each interview lasted about 20 minutes.

Quantitative Analysis

We described a marker-based procedure [5-9] for the quantitative examinations of the full body motion. We describe the different phases involved in the measurement process, including the experimental protocol of acquisition, the human bodies modelling and the data analysis

Human modelling

The stick model of the full-body consists of fifteen anatomical segments: the pelvis, the trunk, the head, the left arm, forearm and hand and the right arm, forearm and hand, the left upper leg, lower leg and the foot, the right upper leg, lower leg and the foot. The segments are connected by fourteen ideal spherical joints. All the joints are characterized by three rotational degrees of freedom, while translation are ignored.

A set of 41 markers, placed in both anatomical and technical repere points, and anthropometrical measurements were used to calculate the position of the joints’ centre. The definition of local reference system for each joints allows to compute the joints angle according to the Eulero-Cardano convention.

Experimental set-up

Fifteen subjects (selected from a clustered population defined by age and mobility skills) were tested. All the considered subjects are skilled into the product utilization in their daily activities.

Each subject were asked to perform a sequentially randomised interaction activity with the fridge starting from different position and grasping three objects with different weight. We gave no instructions about the speed, an we let them move as naturally as they could.

The movements were acquired with a VICON optoelectronic system with 6 infrared M cameras, working at 120 Hz sampling frequency.

A set of quantitative variables was calculated for the ergonomic evaluation of human movements:

- *Joints trajectories.*
- *Joints angles (*to find the motor synergies).
- *Specific Indexes*:
    - *Trajectory Index*: it is defined as the relation between the length of the trajectory of the end effector and the target distance.
    - *Variability Index*: this index quantify the relation between the variability of the trajectories of the joints and their displacement.
    - *Time- Angles Index:* it is defined as the sum of the values of the joint's angle assumed during a path of movement multiplied for the duration of the single execution.
    - *Weight-Angles Index*: it is defined as the sum of the values of the joints' angle multiplied for the proportional weight coefficient according to anthropometrics data from literature.

Data Integration

The aim of this measurements methods was to correlate these different kinds of information and to suggest a new protocol introducing the most significant ergonomic variables since the phase of definition. Results from ethnographic analysis and from the quantitative analysis would provide an ergonomic evaluation of the industrial product.

The sets of qualitative and quantitative ergonomics variables were analysed. The integration process was developed in three phases:

1. The ethnographic analysis provided a set of motor strategy describing the home appliance-human interaction in the daily conditions. This kind of information allowed to focus the object of the quantitative analysis.
2. The optoelectronic system provided an ergonomic evaluation of the observed strategy repeated in different boundary condition. This kind of information allowed to define the most significant variables classification, according to the traditional biomechanical analysis, and that provided a comparison with the qualitative analysis
3. The choice of the observed motor strategy (integrated with the quantitative evaluation in different conditions) allowed the suggestion of new design specification for the project of a more ergonomic design of the home appliance.

Conclusion

The presented method wants to provide a qualitative and quantitative measurement procedure for ergonomics. This method can be applied both for the ergonomic evaluation of an existing product and for the definition of design specification before its realization through mock up analysis. The qualitative analysis identifies the most significant motor strategy in the daily use of the home appliance, while the quantitative analysis provides a set of quantitative variables for the ergonomic classification of the observed strategy.

References

1. Button, G. e W. Sharrock, (1995), Entnomethodology and HCI, in *CHI 95 Research Symposium*, Denver, May 1995 Corbetta
2. Allard P, Cappozzo L, Lundberg A, Vaughan CL (1997), Three dimensional analysis of human locomotion, John Wiley & Sons Ltd.
3. Beyer, H. e K. Holtzblatt (1997), Contextual Design: A Customer-Centered Approach to Systems Designs, Morgan Kaufmann
4. Lifchez, R. e B. Winslow (1979), Design for indipendent living: the environment and physically disabled people, Architectural Press
5. Davis, RB III, Ounpuu, S, Tyburski, D, and Gage, JR (1991). A gait data collection and reduction technique. *Human Movement Sciences* **10**, 575-587
6. Kadaba MP, Ramakristan HK, Wooten ME (1990) "Measurements of lower extremity kinematics during level walking". *Journal of Orthopaedic research* **8**, 383-392
7. Vaughan CL, Davis BL, O'Connor J (1999) Dynamics of Human Gait (Second Edition), Kiboho Publisher, Cape Town, South Africa
8. Kadaba M P, et al. (1989). Repeatability of Kinematic, Kinetic and Electromyographic Data in Normal Adult Gait. *Journal of Orthopaedic Research* **7**, 6, 849-860.
9. Tabakin D. Vaughan C.L. "A comparison of 3D gait models based on the Helen hayes marker set". Department of Biomedical Engineering University of Cape Town, South Africa.

uLog:
Towards attention aware user-system interactions measurement

A. van Drunen[1,2], E.L. van den Broek[2], and T. Heffelaar[1]
[1]*Noldus Information Technology, Wageningen, The Netherlands (a.vandrunen@noldus.nl, t.heffelaar@noldus.nl)*
[2]*Center for Telematics and Information Technology, University of Twente, Enschede, The Netherlands avandrunen@gmail.com, vandenbroek@acm.org*

In user interface development, it is essential to find the balance between grasping and, subsequently, holding attention and providing an optimal information density both in the center and periphery of focus, refraining from information overload [2,7]. This process is like balancing on a thin line: the exact amount and optimal manner of information presentation are crucial [6]. In this paper, we provide an outline on how user interface development can be facilitated, using the program uLog, developed by Noldus IT.

uLog was developed to record user system interaction and, subsequently, determine programs' usability. In addition, we aim to determine user's attention to parts of a user interface (e.g., a website). For example, the latter is of great importance for advertising. uLog records interactions such as keystrokes, mouse clicks, application activations, and mouse movements. We expect that such information can also reveal the pattern of user attention. To verify and validate the uLog recordings for user attention analysis, a research line is developed as will be described in the current paper.

The uLog recordings need to be accompanied by other measures that can reveal user's patterns of attention. The following, complementary data sources were collected in parallel to the uLog recordings:

- Questionnaires, including demographic information
- Eye tracking: fixation points and eye gazes during website usage
- Think aloud of the participants during the task
- ECG recording: determination of Heart Rate Variability of participants
- Skin Conductance (SC) determination during the experiment

The questionnaires, as conducted at both the begin and the end of the experiment, were used to gather demographic data and to determine user's self reported mood, emotions, and evaluation of the stimuli.

Using the think aloud technique, users were asked to verbalize their thoughts during the experiment. In the most commonly used approach, concurrent think aloud (CTA), users' verbalization takes place simultaneously with their task performance [1]. CTA is widely used in research on cognitive processes and for HCI processes. CTA data provides a good indication on users' attention while performing certain tasks.

The eye tracker is integrated in the computer screen and detects the gazes, which are called saccades, and fixation points of the users' eyes. When exploring an object or a website or when reading a text, users typically show a pattern of a saccade followed by a fixation [4]. These patterns of viewing behavior are recorded through the eye tracker and, subsequently, can be visualized, providing information on the attention patterns of the user.

Heart Rate Variability (HRV) is one of the physiological measures recorded in the experiment and refers to the variability of the heart period over time or between beats [8]. Research has resulted on contradicting results with respect to the correlation between user's HRV and user's mental state. Though, some evidence is present for a negative correlation between mental load and HRV [9]. Therefore, HRV is included as a measure, in particular to determine the amount of workload and attention experienced by users.

As a second psychophysiological measure, Skin Conductance (SC) [8,10] is included. Differences in SC also indicate possible changes in attention or workload [3]. As for HRV, various factors (e.g., stress and emotion) can heavily influence the results of this measure. Hence, these factors have to be controlled as much as possible.

The variety of measures that are recorded in parallel with the uLog recordings provide a multiperspective view on users' behavior, in particular their pattern of attention. These measures are incorporated in one set of analyses, as the principle of triangulation proposes. Triangulation refers to use of multiple sources and/or analyses to capture a construct. This approach reduces the amount of error in interpretation of users' behavior. Errors can be detected more easily and isolated from the construct itself [5]. Especially when taking in consideration the complexity of the phenomenon of attention, triangulation is needed to derive conclusions from the measures recorded. Possibly, not all measurements are needed since they are not completely complementary to each other. The current research outline will reveal this also. Most important it can provide us solid evidence on the patterns of attention of users and can show us how they relate to uLog recordings.

Although the current research outline can be considered as rather obtrusive and, therefore, unwanted in practice, it can provide us the way to unobtrusive recording of users' patterns of attention, using the uLog recordings. uLog can record user-system interaction even without awareness of the user and without any disturbance or delay in the interaction process. However, even when solely uLog data would not provide enough information, a selection of other measures can be made that, on the one hand, optimally complement uLog recordings and, on the other hand, limit the obtrusiveness of the recordings for the user.

Taking all in consideration, a unique research outline is introduced that can be an important step towards automated analysis of users' behavior and of their patterns of attention.

References

1. Boren, T., Ramey, J. (2000). Thinking aloud :reconciling theory and practice. *IEEE transactions on professional communication* **43(3)**, 261-278.
2. Buscher, G. (2007). Attention-Based Information Retrieval. *Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval, 918*
3. Critchley, H. D (2002). Electrodermal responses: What Happens in the Brain. *The Neuroscientist* **8(2)**, 132-142.

4. Duchowski, A. T. (2002). A Breadth-First Survey of Eye Tracking Applications. *Behavior Research Methods, Instruments, & Computers* **33(4)**, 1-16.
5. Heath, L. (2001). Triangulation: Methodology. In N. Smelser & P. Baltes (Eds.), *International Encyclopedia of the Social and Behavioral Sciences (Vol. 23)*, 15901-15906.
6. Horvitz, E., Kadie, C. M., Paek, T., Hovel. D. (2003). Models of Attention in Computing and Communications: From Principles to Applications. *Communications of the ACM* **46(3)**, 52-59.
7. Johnson, A., Proctor, R.W. (2004). Attention: Theory and practice. London, UK: Sage.
8. Stern, R. M., Ray, W. J., Quigley, K.S. (2001). Psychophysiological Recording. Chicago, USA: Oxford University Press.
9. Wastell, D.G., Newman, M. (1996). Stress, control and computer system design: A psychophysiological field study. *Behaviour and Information Technology* **15(3)**, 183–192.
10. Westerink, J. H. D. M., Broek, E. L. van den, Schut, M. H., Herk, J. van, Tuinenbreijer, K. (2008). Computing emotion awareness through galvanic skin response and facial electromyography. In J. H. D. M. Westerink, M. Ouwerkerk, T. Overbeek, W. F. Pasveer, B. de Ruyter (Eds.) *Probing Experiences: From Academic Research to Commercial Propositions (Part II: Probing in order to feed back, Chapter 14; p. 137-150)*. Springer: Dordrecht, The Netherlands.

Measuring computer use behavior: patterns of variability, within and across days and between different users

H.P. Slijper[1], S.E. Mathiassen[2] , E.A.B. Over[1], J. Richter[1], and M.A. Frens[1]

[1] *Department of Neuroscience, Erasmus MC, Rotterdam, The Netherlands, harm.slijper@gmail.com*

[2] *Centre for Musculoskeletal Research, University of Gävle, Sweden, svenderik.mathiassen@hig.se*

Introduction

Computer users are notoriously bad at estimating how much time they have spend using the computer. Levels of overestimation are large (up to 120%) and depend only slightly on the amount of time actually worked with the computer [1]. Thus, in order to arrive at an objective and precise description of computer use behavior we developed registration software which unobtrusively recorded input device use in a large group (N=500) of heavy computer users. The computer users were followed up to several years to characterize the **temporal** (pause behavior) and **spatial** (mouse use) patterns that governed their computer use behavior. With regard to the temporal pattern, we described the episodic nature of work and non-work episodes both *within* days and *across* days. We asked the following questions: 1. How do work/ pause patterns change when the computer is more intensely used? 2. Can the characteristics of a working day be predicted on the basis of recordings from previous days? 3. How well can the average computer use (across a year) be predicted from recorded computer use during only a few days (this determines ergonomist's measurement strategies)? 4. To what extend is computer usage subject specific? With regard to this last question we also looked at the spatial pattern of mouse movements to see whether the characteristics of mouse movements varied systematically between users.

Methods

The recorded time traces of computer use (N>72.000) included timestamps of keyboard strokes and cursor changes (10 Hz). In such a time line of events, which have no duration themselves, "computer-work" and "non-computer work" episodes need to be discriminated. To this end, we implemented a temporal criterion (non computer threshold, NCT) that specifies the amount of time two subsequent events could be separated in time, while the time in between would still be classified as continuous work [2]. On the basis of this NCT, the duration of the working day, the summed duration of the work and non-work episodes and the number of work episodes, were used to describe the pattern of computer use within a work day. By varying the NCT value we investigated how the values of these variables systematically changed across different time scales (i.e. a small NCT will identify a lot of short duration episodes, while a large NCT will identify few long duration episodes). The shape of the relationship between NCT will thus be able to detect whether there is regularity in the usage behavior. To analyze whether the characteristics of a workday could be predicted based on previously recorded days, we calculated day-to-day autocorrelation functions for the above variables. To calculate the error in estimating "the one-year-average duration of computer use" we used empirical re-sampling of data. That is, by randomly drawing subsets (certain number of sample days) of data from all available data from a subject, and repeating this procedure thousands of times (bootstrap-like procedure), we could calculate how reliable the mean work duration across a work year could be estimated. For all variables, we performed a variance component analysis to see how much variability could be explained by differences between subjects and how much by differences between work days. In a separate analysis, we studied the kinematics of mouse movements [4]. Using a velocity threshold, begin and endpoint of individual mouse movements were identified. Consequently, we focused on characterizing directional distributions, since there was a strong bias of users to make movements in particular directions. We investigated whether these directional distributions were invariant across days and computers (used by the same worker).

Results

Q1: We found that distributions of episode (durations) were highly skewed (al lot of small episodes and few large ones). These distributions depend on the chosen NCT according to a log-linear relationship [3] (figure 1). While the time classified as work increases with a mere 3.5% per doubling of the NCT, the number of computer work episodes decreased by 40%, the duration of computer work episodes increased by 90% and the duration of non-computer work episodes increased by 60%. This means that when the duration of a non-work episode is doubled it is 1.6 times less likely to occur. The slopes of these relationships remained invariant while the intercept changed when workers used the computer more intensely (for example see Figure 1).

Q2: Since the day-to-day autocorrelations (r = 0.1-0.2; non significant) were low across a large number of lags (days) and all variables, the characteristics of a working day could not be

Figure 2: Error in time working with the computer for different sample sizes

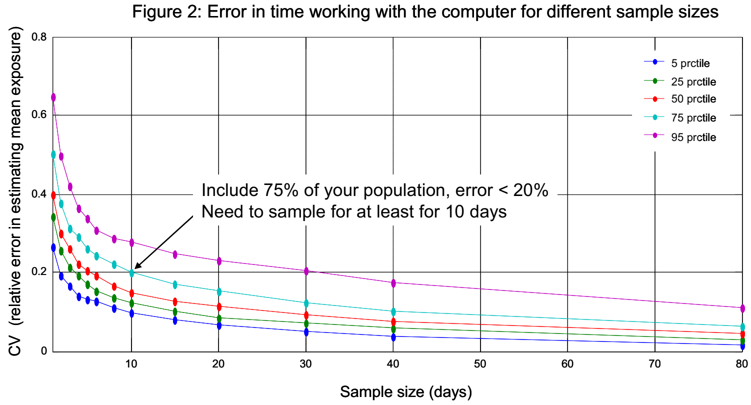


Figure 1: time classified as work depending on NCT

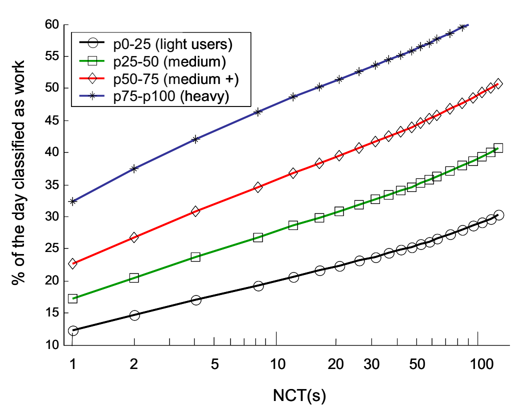
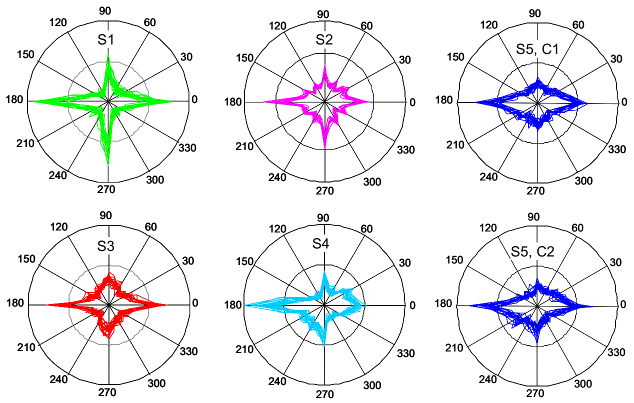


Figure 3: directional distributions in 5 users across 25 days



predicted precisely from the recordings of previous days.

Q3: In order to estimate the reliable assessment of mean computer use duration across a 1-year period, we calculated relative errors (CV) for different sample sizes, as shown in figure 2. This shows that considerable days need to be sampled in order to estimate 1-year average values reliably.

Q4: Variance component analysis showed that between 20-50% of the variance could be explained by differences between workers, depending on the variable. A striking example of a subject specific pattern was found in the directional distribution of mouse movements (Figure 3). Shown is data from 5 users (S1-5), across 25 workdays (different lines). User 5 worked on two computers (C1, C2). Note the preference for movements in cardinal directions, the invariance of the distribution across days and computers, and the idiosyncratic differences between users.

Conclusions

Although computer users are free to choose the onset and duration of their work episodes, the (spontaneous) pattern of work and non-work episodes is highly structured Day-to day variability in computer use is unpredictable which necessitates long recordings to estimate 1-year averages (unfortunate for ergonomists). Subject-specific patterns in the directional distribution of mouse use show that users have a 'mouse-signature'.

References

1. Faucett, J., Rempel, D., 1996. Musculoskeletal symptoms related to video display terminal use: an analysis of objective and subjective exposure estimates. Aaohn J, **44(1)**: 33-9.
2. Slijper, H. P., J. M. Richter, et al. The effects of pause software on the temporal characteristics of computer use. *Ergonomics*, 2007, **50(2)**: 178-91.
3. Richter JM, Slijper HP, Over EAB, Frens MA. Computer work duration and its dependence on the used pause definition. *Applied ergonomics*. 2008, **Jan 3** (online first).
4. Slijper, H.P., Richter, J.M., Over, E.A.B., Smeets, J.B.J., Frens, M.A (2008) Statistics predict Kinematics of Hand Movements during Everyday Activity, *Journal of Motor Behavior* (in press).

Evaluating a location-based mobile game in early stages of the development

M. Haesen, C. Raymaekers, and K. Coninx
Hasselt University - transnationale Universiteit Limburg, Expertise Centre for Digital Media - IBBT, Diepenbeek, Belgium, {Mieke.Haesen, Chris. Raymaekers, Karin.Coninx}@uhasselt.be

Introduction

When employing a User-Centered Design (UCD) approach for the development of an interactive system, the active involvement of end-users is recommended [2]. Within the VIP-lab project [1], some case studies were carried out using UCD in five different application domains. This contribution will introduce two techniques to measure the user's experience for a location-based mobile game.

Location-based mobile game

Since children seem rather uninterested when they visit a museum or a park, the aim of the case study for the domain of tourism and culture was to make school excursions more interesting for children between the age of 8 and 10. During first meetings with cultural and tourist organizations, it became clear that a prototype of a mobile location-based application might be suitable to attract the children's attention, but only if it would offer more than just an informative guide on paper.

In order to get an overview of the needs of the end-users, researchers joined several school groups visiting a museum or a park, and observed the children. Based on this user and task analysis, a list with points of interest was created. For instance, children liked the tours that were based on a story providing some excitement and suspense.

These points of interest were used to create two game concepts within a multi-disciplinary team. Besides computer scientists and social scientists, a graphic designer and delegates of a mine museum and a nature resort attended brainstorm sessions.

Adding locations to a lab

To measure the user experience of the game concepts, some first designs of the user interface were evaluated. The fidelity of the prototype implied that the first evaluation took place in a lab. Since the test persons should visit several locations in a nature resort and a mine museum while playing the games, two physical locations in the lab were used. On these physical locations beamers projected virtual locations on white screens (see figure 1). Using materials similar to what is used in participatory approaches such as PICTIVE [3], the behavior of the game was simulated. Screenshots for the two games, designed by the graphic designer, were printed on paper and could be stuck on a plasticized picture of a PDA that the test users carried with them while they walked from one location to another.

Four groups of two children tested the games, while two researchers were observing, and two other researchers simulated the role of the computer by providing the audio of the game, sticking the correct screens on the picture of the PDA and making available the projections of virtual locations. Although we used basic materials instead of an interactive prototype, the test users were very amused while playing the games and loved the designs. The concept of walking to several locations became obvious to the children after some hints of the facilitator, but they sometimes forgot to take the picture of the PDA with them. In a later stage of the case study we experimented with a virtual model of the mine museum to add locations to a lab environment. First results showed that walking around in this virtual environment was more intuitive for the children.

The test results revealed a few weaknesses in the user interface, but showed on the other hand that the game concept was clear for children.

Involving a wizard in field tests

The next iteration of the case study concerned the development of an interactive prototype for a PDA and its evaluation in the field. Since it was impossible to use a similar location detecting technology for both the indoor mine museum and the outdoor nature resort, a Wizard of Oz application was developed, which was deployed on a second PDA, connected to the game PDA using an ad-hoc Wi-Fi connection.

Using the Wizard of Oz application, location information could be send to the game to help the children finding their way, or to trigger a new task. During field tests in the nature resort and the mine museum, ten groups of two to four children played the game. Simultaneously a researcher (wizard) operated the Wizard of Oz application (see figure 2). Although the wizard was participating to the test, the test users did not notice that the location information was passed on by the Wizard of Oz application. Afterwards children were enthusiastic about the game and they even asked if this prototype can be used to play location-based games at home.

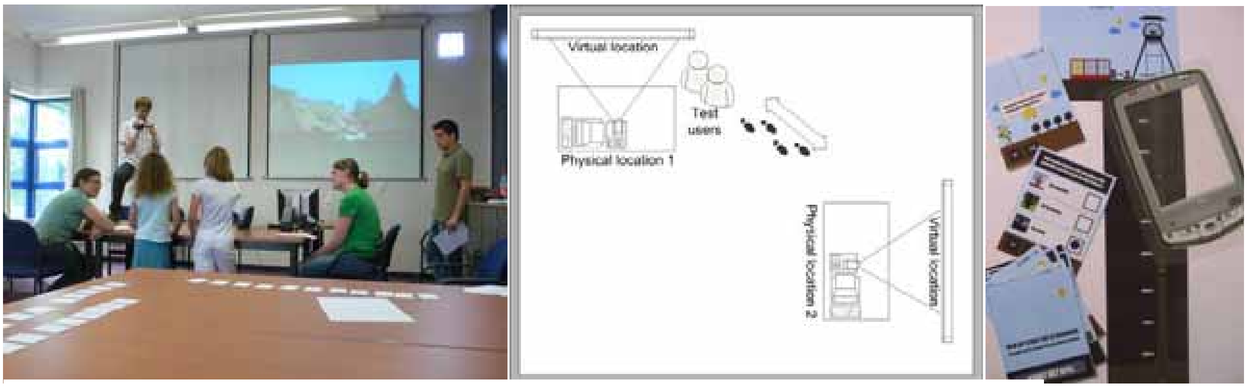


Figure 1. *The setup and materials for the lab test of the location-based mobile game.*

Figure 1. *The "wizard" sends location information to the game PDA.*

Conclusions

In this contribution we presented two techniques to evaluate a location-based mobile game when detailed location information is missing. By evaluating after each iteration, in the lab or in the field, the prototype can be fine-tuned in a cost-effective way. These techniques provide suitable results and can be used to measure the user experience of location-based prototypes in early stages of the development.

Acknowledgements

VIP-lab is financed by the "Euregio Benelux-Middengebied" authorities and is co-financed by Province of Limburg (B), Province of Limburg (NL), Ministry of Economic Affairs (NL) and Ministry of Flemish Government/Economic Affairs (B). We acknowledge the contribution of the consortium.Part of the research at EDM is funded by EFRO (European Fund for Regional Development), the Flemish Government and the Flemish Interdisciplinary institute for Broadband technology (IBBT).

References

1. Coninx, K., Haesen, M. and Bierhoff, J. (2005). VIP-lab: A virtual lab for ICT experience prototyping. *Proceedings of Measuring Behavior 2005 (Wageningen, 30 August - 2 September 2005)*, 585-586
2. International Standards Organization (1999). ISO 13407:1999(E) – Human-centered design processes for interactive systems.
3. Muller, M.J. (1991). PICTIVE - An exploration in participatory design. *CHI '91: Proceedings of the SIGCHI conference on Human factors in computing systems (1991)*, 225-231

Epistemic action: A measure for cognitive support in Tangible User Interfaces?

M. Fjeld[1] and W. Barendregt[2]

[1]*TableTop Interaction Lab (www.t2i.se), CSE, Chalmers TH, Göteborg, Sweden, morten@fjeld.ch*

[2]*IT University, University of Gothenburg, Göteborg, Sweden, wolmet@ituniv.se*

Introduction

In the field of Human-Computer Interaction the usability of a program or tool is often measured in terms of efficiency, effectiveness, and satisfaction [1]. Recently, in the field of Tangible User Interfaces (TUIs) an additional measure for the quality of tangible tools was introduced: epistemic action. Kirsh and Maglio [2] distinguish between 'epistemic' and 'pragmatic' actions. An epistemic action is an action whereby users change their environment to search for a solution or strategy to perform a certain task. A pragmatic action is strictly the action needed to actually perform this task. Kirsh and Maglio illustrate this with the example of how players of the game, Tetris, rapidly rotate the falling bricks instead of mentally determining the correct position for a brick and then rotating it to the correct position. Players use epistemic actions to modify the environment which helps them to determine the correct position. They can do this faster than the corresponding mental rotations. It might be easier to physically modify the external world and then interpret it rather than compute and interpret a new state internally. It has been suggested that epistemic action is a relevant concept when researching computer interfaces that involve physical objects such as TUIs [3]. Sharlin et al. [4] suggest that support for epistemic actions is an important factor in the success of a TUI. In this paper we look at three different spatial planning tools and aim to relate the traditional measures of efficiency, effectiveness, and satisfaction with the number of epistemic actions to determine whether this last measure can additionally be a useful measure of quality for TUIs.

Epistemic action and TUIs

According to Fitzmaurice [5] epistemic actions can support a user's cognition by:

- Reducing the memory involved in mental computation
- Reducing the number of steps in mental computation
- Reducing the probability of error of mental computation

TUIs offering tangible objects that can be physically manipulated may offer more cognitive support than interfaces without these objects because they support epistemic actions.

In order to study epistemic action as a quality measure we designed a spatial planning task [6]. The task was to find the unique of many alternative blocks on which to place a laser source in order to hit a nearby target with its light beam. The beam should be as close as possible to the centre of the target. Nine square blocks with different heights formed a three-by-three matrix. The laser source could be placed and slightly adjusted on any of the nine blocks.

We chose three different tools to fulfil this task, each offering a different degree of physical interaction:

1. No physical interaction: Modeller, a CAD tool with virtual tools and views [7].
2. Some physical interaction: BUILD-IT, which employs a virtual modelling of the blocks, target, and laser source. The participants use one physical brick to manipulate the virtual laser source.
3. Only physical interaction: PhysicalBlocks, consisting of nine metal blocks, a standard laser source, and a target consisting of a metal pin attached to a metal flag. The participant can adjust the height of the metal flag, the target position, and the block positions.

Experiment and results

In an in-between-subject experiment [8] we measured efficiency (trial time), effectiveness (percentage of correct trials), satisfaction (questionnaire), and epistemic action (average number of tested blocks in a trial). Each tool was assigned to ten participants who each had to perform ten task variations. PhysicalBlocks yielded the lowest trial time, the highest percentage of correct trials and the highest user satisfaction. Modeller yielded the highest trial time, lowest percentage of correct trials, and the lowest user satisfaction. BUILD-IT yielded results in the middle of the other two for all measures. These results can be related to how much physical interaction each tool offers. As for epistemic action, we saw a result that was not related to any of the traditional usability measures. The number of tested blocks in a trial was lowest for PhysicalBlocks, but highest for BUILD-IT and in the middle for Modeller. Epistemic action measured by the number of blocks tested was not directly related to the level of physical interaction offered by the tool. The tool that offered most physical interaction was, indeed, the tool with the lowest number of blocks tested in a trial.

Discussion and conclusions

Our experiment makes us rethink epistemic action as a simple linear measure for the successfulness of a TUI. A primary observation is that PhysicalBlocks offers users the possibility to change the position of their head to determine the correct block. This will not result in a countable epistemic action expressed by testing a block. Furthermore, it is also possible that PhysicalBlocks offers so much support in the physical world that epistemic actions to modify the world are not even necessary. Since epistemic action may take different forms, using it as a measure for TUI success should be considered with care. In future work, we firstly need to determine more specifically what should be considered an epistemic action. Secondly, it is possible that increased possibilities for physical interaction decrease the need for modifications in the environment in order to find a solution or strategy. This might result in a lower number of epistemic actions for tools that provide more cognitive support.

References

1. ISO 9241-11. Ergonomic requirements for office work with visual display terminals (VDTs) - Part 11: Guidance on usability (1998).
2. Kirsh, D., Maglio, P. (1994). On Distinguishing Epistemic from Pragmatic Action, *Cognitive Science*, **18**, 513-549.
3. Patten, J. and Ishii, H. (2007). Mechanical constraints as computational constraints in tabletop tangible interfaces.

Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (San Jose, California, USA, April 28 - May 03, 2007). CHI '07. ACM Press, New York, NY, 809-818.

4. Sharlin, E., Watson, B.A., Kitamura, Y., Kishino, F., and Itoh, Y. (2004). On humans, spatiality and tangible user interfaces. *Pervasive and Ubiquitous Computing, Theme issue on Tangible Interfaces in Perspective 8*, 338-346.
5. Fitzmaurice, G. (1996). Graspable User Interfaces. Ph.D. Thesis. Department of Computer Science, University of Toronto.
6. Fjeld, M., Guttormsen Schär, S., Signorello, D., Krueger, H. (2002). Alternative Tools for Tangible Interaction: A Usability Evaluation. In *Proceedings of the IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR 2002)*, 157-166 .
7. Gähwiler, M., Fjeld, M. (2002). Evaluation of CAD-systems as alternative tools to a Tangible User Interface (TUI). ETH, Eidgenössische Technische Hochschule Zürich, IHA, http://e-collection.ethbib.ethz.ch/show?type=semarb&nr=44
8. Wirz, R. (2005). Evaluation of a comparative study of a tangible user interfaces. http://www.t2i.se/pub/papers/Raphael_Wirz_2005.pdf

Measuring design behavior: Analysis of networks of links among speech units in design sessions

G. Goldschmidt
Faculty of Architecture and Town Planning, Technion – Israel Institute of Technology, Haifa, Israel, gabig@tx.technion.ac.il

Capturing Design Behavior

Design behavior usually refers to designers' activities while engaged in work towards producing a design proposal. This may mean team or individual work, at any phase of the design process. Capturing and documenting design behavior necessitates evidence from the process itself, i.e. design sessions, which can be collected by observation and recording (audio/video), and post hoc interviews with participants. Data thus collected, along with by-products of design sessions (sketches, notes) can be submitted to analysis. In the design thinking research literature a prevalent research methodology is what might be by now termed 'classic' protocol analysis, along the lines proposed by Ericsson and Simon more than two decades ago [1]. Many of the studies undertaken pertain to short-term design sessions (not exceeding two hours), in the early stages of conceptual design. In 'classic' protocol analysis verbalizations are analyzed: they are recorded and transcribed into a protocol which is then parsed into speech units. Those units which may be short or somewhat longer – depending on the nature of the study – are then coded, using schemes of categories that befit the aims of the study. Various statistical tools may be used to analyze the quantitative results the coding yields. In a design session in which an individual designer works alone, verbalization is prompted by asking the person to think aloud. Cross et al. [2] published a collection of protocol analysis studies using protocol analysis variants, all pertaining to the same two design protocols.

Linkography

Linkography was developed as a notation and analysis system based on protocols, see e.g., [3]. Unlike 'classic' protocol analysis, it concerns itself not with coding but with links among the speech units into which a protocol is parsed (by extension, it may also deal with other units or longer sequences of speech. For example, ideas extracted from a protocol). In many studies the unit of speech is a design move – the smallest step made by a designer, as a result of which the state of the design situation is at least somewhat transformed. In conceptual design the duration of a move may be a few seconds and the analysis is accordingly at the cognitive level. Since design problems are generally ill-defined and ill-structured, protocols of conceptual design phases reflect a search for a solution in which moves cannot be predicted ahead of time. Links among speech units (moves for the most part) therefore reflect the process of reasoning the designer is engaged in, as he or she endeavors to synthesize the primary concepts of a design solution.

Links between every pair of speech units are notated in a Linkograph, which is a diagram derived from a matrix. To do so, every unit is paired with every one of the preceding units and for every pair we ask: is there a link between the two units? A link pertains to the contents of the units, and is established using (expert) common sense. Thus for unit n the question is asked n-1 times, to check for links between n and 1, 2, 3… n-1. The total number of checks, and therefore of potential links is n*(n-1)/2. The system is binary in the sense that only a yes or no answer is given. Further coding of links is possible but no mandatory. Because of the large number of checks the process is labor intensive and therefore this method is suitable only for relatively short sequences of speech units. A typical Linkograph is shown in Figure 1, where links are represented by dots in a network.

Linkographs allow us to visualize the pattern of links among speech units and measure its properties in the form of a number of variables. The main variables are:
Link index: the ratio of links per speech units/moves.

The number and proportion of critical speech units/moves. A critical unit is one with a relatively high number of links associated to it; the threshold number of links used to determine criticality is arbitrary and is established per study relative to the grain of the study. We distinguish between criticality due to backlinks and criticality due to forelinks. Backlinks are links between a unit and previous units: in generating a Linkograph only backlinks are established. Forelinks are the virtual links between a unit and subsequent units. Such links can only be established post factum; in a Linkograph they have the same status as backlinks. Every link is counted once as a backlink of a particular unit, and once as a forelink of the other unit in the pair of linked units. The sum of all back and fore links is therefore twice the total number of links in the Linkograph.

Other variables pertaining to link distribution and positioning are also measured in Linkography and attest to the structure of design reasoning..

Linkography is particularly useful in comparisons, e.g., between processes by different designers in the same task; different phases in the same process; different predefined groups of designers, for instance experts and novices. It has also been used to compare communication parameters such as interactivity in tutorial conversations between teachers and students in studio sessions; the ratio of idea generation by students in a studio setting; creativity in idea-generation sessions of design teams; and the effect of different sources of inspiration on creativity and fixation. Linkography variables have been correlated with design productivity and creativity and appear to be useful in micro studies of cognitive behavior and reasoning processes in the context of design and design education. We presume that this method has the potential of illuminating matters of verbal communication in general, particularly in creative problem solving.

References

1. Ericsson, K.A., H.A. Simon, H.A. (1984/1993). *Protocol analysis: Verbal reports as data*. Cambridge, MA: MIT Press.
2. Cross N, Christiaans H, Dorst K. (eds.) (1996). *Analyzing design activity*. Chichester: Wiley.
3. Goldschmidt, G. (2003). Cognitive Economy in Design Reasoning, in Lindemann, U. (ed.), *Human Behaviour in Design*. Berlin: Springer Verlag, 53-62.

How User Oriented Methodologies can help to build better services?

R. Lahuerta[1], M.J. Such[2], C.Soler[3], P. Vera[4], and S. Redondo[5]
[1]*Institute of Biomechanics of Valencia, Valencia, Spain, ruben.lahuerta@ibv.upv.es*
[2]*Institute of Biomechanics of Valencia, Valencia, Spain, mariajose.such@ibv.upv.es*
[3]*Institute of Biomechanics of Valencia, Valencia, Spain, carlos.soler@ibv.upv.es*
[4]*Institute of Biomechanics of Valencia, Valencia, Spain, pedro.vera@ibv.upv.es*
[5]*ER&SI, Valencia, Spain, info@er-si.com*

Design strategy involves not only a product conception and development, but also its launching and commercialization stage. Is a product's commercial success related to the context of the sale? Does store environment encourage or discourage customers? Could purchase experience be designed?

The User Oriented Development model (strategic definition, design, validation of prototypes, and commercialization/ launch) offers multitude of applicable tools and methodologies.

It is in the phases of strategic definition, design and validation where the majority of these methodologies and tools are employed. The promotion and sale is the phase in which these techniques and tools are less employed for customization according to the user voice.

An adaptation of the UOD methodology to the commercialization and launch phase, will allow to really expand the UOD techniques to this phase and to develop a customization service methodology to employ in the phase of promotion and sale of new products.

The customization service in the phase of commercialization and launch of new products provides an opportunity to increase the success of the UOD model.

In order to shed light on these questions a theoretical and practical research has been made, to generate, or adapt tools and methodologies, that enable the obtainment of valid data that helps in the design of user purchase experiences. A set of UOD (User-Oriented Development) methodologies were used in two research studies:

1. When the situation implies choosing a product that is a part of a whole, the user is faced with the effort of abstracting in order to "see" how it is going to fit for its purpose. According to some researchers [8], the purchase experience has an implicit stress component associated to the decision making process. As the selection of products to choose from increases, the visualization of the outcomes gets more and more complex, therefore the possibility of consumers negative affect raises [6]. In spite of how trendy designed the point of sale is, this does not guarantees that it will meet with the user's image of an ideal setting, thus affecting his or hers purchase intention.

On the other hand, several authors point out that the measurement of consumer satisfaction is a weak indicator regarding consumer loyalty or willingness to pay [1] [3], so measurement of other indicators is necessary. It is highly accepted that consumption is associated with hedonic pursuits such as fun and pleasure [4] and that every product carries a symbolic meaning [7], therefore to design a retail environment, meanings and emotional implications of the consumer are essential.

The EMOCIONA initiative has demonstrated that the application of Emotional Design Techniques can serve to improve the design of surrounding settings used in the point of sale of habitat-related products, and to measure to which extent a retail store's background has influence in the willingness to purchase a product. Around eighty people had participated of this experience through a pilot emotions measurement laboratory in which the emotional profile of the users was determined and its purchase attitude was registered, in order to extract concepts associated not only with the piece being evaluated, but also to establish the influence that different scenarios had in their perception of it.

The results showed an increase of the purchase intention of a piece of furniture when it was settled in users favourite environment, and also the emotional component the "desired environment" concept responds to in terms of style, global preferences, et cetera.

Furthermore, the results had permitted the corroboration of the possibilities that the utilisation of Emotional Engineering techniques, as part of User Oriented Development methodologies, have in the area of personalization of a retail store environment; allowing its adaptation for a determined population and the effective communication of a brand image.

2. In the particular case of personalized PAL´s (Progressive Addition Lenses) we pretended to go further. An extensive research was made in order to find out the current situation of the eyecare selling business about both; technology and marketing techniques. It was established that there is a growing trend to use dispensing tools to improve the user's purchase experience [2] [9] and that this is setting a baseline for what users expect from a service.

Research has shown that the devices are publicized as tools for reaching different kinds of personalization such as; 1) Lens and Frame fitting systems; 2) Lens and Frame demonstration devices; 3) Lifestyle profiling systems and 4) Visual behavior diagnostic devices [9]. EyeMade personalized progressive addition glasses, are designed according to the user's VisualMap which is obtained through the Visual Map Developer. Since EyeMade lenses are a highly personalized item, which uses a state of the art diagnostic technology for its individualized design, it falls within the limits of the dispensing tools used in the sector. Subsequently existing experiences and studies of the influence of this technology where used as a basis to elaborate a theory on what a user may expect from the EyeMade experience.

According to several researchers [10], the customer formulates his/hers expectations according to a confirmation/disconfirmation mental model. As such, when a service does not meet the expectations a mismatch occurs; if a service matches or exceeds expectations is perceived as a positive experience but, if it is below what is expected the experience is perceived as negative. A theory on value [3] points out that costumer willingness to pay increases when a unique service is delivered to them [5], hence this creates a differentiation value that is the costumer perception of the favourable uniqueness of a product/service among other similar ones. Whereas merchandise value implies that good quality is delivered at a low price, differentiation value occurs when the highest quality, tangible or not, is offered at a higher price [3].

As the Visual Map Developer allows a highly grade of individualization of the product, - visual strategy is individual and different for every user [7]-, it has an intrinsic, and very valuable, differentiation factor. This is at the same time, a powerful selling point and a fine service for the customer. Furthermore if we can tailor the service to the clients needs and expectations (or exceed them), then we will have a highly competitive product.

A series of hypothesis were established related to what the user experience should be like, and which elements of it are important in the decision making process.

This paper pretends to explore the potential that UOD (User-Oriented Development) methodologies have to help build services to generate exceptional purchase experiences.

References.

1. Barlow, J. & Maul, D. (2000). Emotional value: Creating Strong bonds with your customers. San Francisco, CA: Berrett-Koehler.
2. Chasnov, C. (2006). Offer patients Custom Eye Wear. *Optometric Management* **58**, 31.
3. Chaudhuri, A. (2006). Emotion and Reason in Consumer Behavior. Burlington, MA: Elsevier.
4. Holbrook, M. & Elizabeth, H. (1982). The Experiential Aspects of Consumption: Consumer Fantasies, Feelings and Fun. *Journal of Consumer Research* **9**, 132-140.
5. Ligas, M. (2000). People, products and pursuits: Exploring relationship between consumer goals and product meanings. *Psychology and Marketing* **17**, 983-1003.
6. Luce, M. F. (1998). Choosing to avoid: Coping with negatively emotion-laden consumer decisions. *Journal of Consumer Research* **24**, 409-433.

Measuring behaviors of dyssynchronous patient ventilator interaction

K.G. Mellott[1], M.J. Grap[1], and P.A.Wetzel[2]
[1] *School of Nursing, Virginia Commonwealth University, Richmond, Virginia, mellottkg@vcu.edu*
[2] *School of Engineering, Virginia Commonwealth University, Richmond, Virginia*

Background/Significance

Synchronous patient-ventilator interaction is essential in achieving optimal oxygenation and ventilation in critically ill patients. While recent studies show that patient ventilator dyssynchrony (PVD) is highly prevalent in the intensive care unit, it is difficult to quantify. Nurses often assess PVD using a non-empirical approach, and clinical behaviors of dyssynchrony are not standardized. A literature review revealed that PVD is associated with agitation, facial and respiratory behaviors and there is little empirical description of cardiopulmonary changes: tachycardia, tachypnea, decreased oxygen saturation and increased intrinsic positive end expiratory pressure. Describing PVD will assist clinicians in early and accurate detection of dyssynchrony, which may lead to reduction in risks associated with mechanical ventilation and sedation.

Specific Aim

The primary aim of this study is to identify the behavioral and biological markers of PVD.

Methods

A specific list of PVD behaviors was tested and refined during a pilot study. Videography is being used in the current study to capture agitation, facial and respiratory behaviors of PVD for 1.5 hours per subject. A video camera is connected to the computer with The Observer XT 7.0 software installed (Noldus computer) through a firewire cable. The Noldus computer is then connected to the data acquisition system, Biopac (Model MP 150), along with a synchronizing cable to send synchronization signals to the Biopac system. The Biopac software is configured to record the signals in a separate channel from physiological data. Video recorded data are coded for the refined behaviors using The Observer XT (Noldus Information Technology, Wageningen, The Netherlands). The physiological data is later exported from the Biopac system and imported into The Observer XT. It can then be viewed in synch with video after data coding of behavior and airway pressure/flow- time waveforms.

Patient ventilator dysynchrony is identified through physiological signals, ventilator airway pressure- time and flow- time waveforms. These waveforms are obtained through a pneumotachometer connected to the subject’s ventilator circuit (Non-Invasive Cardiac Output Cardiopulmonary Management system [NICO], Respironics, Model 7300, Wallingford, CT) and sent to the Biopac for synchronization. Waveform data are being analyzed for the occurrence of PVD based on Nilsestuen and Hargett, 2005 [1]. Morphological changes in the waveform data are also coded using The Observer XT to represent three major categories of dyssynchrony, trigger, flow and termination. Trigger dyssynchrony has four subcategories: missed inspiratory effort, missed expiratory effort, failure to trigger and double trigger. Termination trigger has two subcategories: premature and delayed termination.

The respiratory markers of physiological instability from PVD are obtained from the NICO using pulse oximetry for oxygen

Figure 2. *Photograph of Biopac cart (right) with NICO and Criticare Scholar in patient room next to ventilator (left) (Photo courtesy of Anne Hamilton, Project Director of SAVE study)*

saturation and the pneumotachometer for respiratory rate and end tidal carbon dioxide. Cardiac physiologic instability is being measured by heart rate obtained from cardiac electrodes placed on the subject and connected to Criticare Scholar III (Criticare Systems, Inc., Waukesha, WI) heart monitor. All cardiopulmonary measures are then sent to the Biopac for time synchronization (See Figure 2 for diagram of Biopac cart).

Results

The Observer XT will allow for successful coding of both videotaped behaviors and physiologic signals of airway pressure and flow. Viewing behavior and the ventilator waveforms simultaneously provides the ability to describe actual subject behaviors of PVD with a standardized measure. It also provides the ability to detect possible time delays from a dyssynchronous event to a real-time behavior change, which has not been described. Therefore, this method will provide the capability to describe the natural history of PVD at a level that has not yet been disseminated Using this method, we will also be able to categorize the behaviors that occurred during different types of dyssynchrony. Coding of the airway pressure- time and flow- time waveforms to describe PVD will provide a convenient application to interface with analysis of subject behavioral data.

References

1. Nilsestuen, J.O. & Hargett, K.D. (2005). Using ventilator graphics to identify patient ventilator asynchrony. Respiratory Care, 50, 202-234.

The project described was supported by Grant Number F31NR009623 from the National Institute Of Nursing Research. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute Of Nursing Research or the National Institutes of Health.

Towards a Standardization in the Use of Physiological Signals for Affective Recognition Systems

J. Arroyo-Palacios and D.M. Romano
Department of Computer Science, University of Sheffield, Sheffield, United Kingdom, j.arroyo@dcs.shef.ac.uk

Abstract
The implementation of physiological signals, as an approach for emotion recognition in computer systems, is not a straight forward task. This paper discusses five main areas that lack of standards and guided principles, which have led Human-Computer Interaction (HCI) researchers to take critical decisions about (i) models, (ii) stimulus, (iii) measures, (iv) features and (v) algorithms with some degree of uncertainty about their results. Methodology standardization would allow comparison of results, reusability of findings and easier integration of the various affective recognition systems created. The background theory is given for each of the five areas and the related work from psychology is briefly reviewed. A comparison table of the HCI common approaches of the five discussed areas is presented, and finally some considerations to take the best decisions are discussed. The aim of this paper is to provide directions on which the future research efforts for affective recognition in HCI should be focused on.

Introduction
Physiological signals provide an insight into human feelings, which are not always completely expressed as facial expressions, body movements, voice tone, and not at all captured in questionnaires. Some of their favourable characteristics as an approach to corroborate affective states in HCI include: reliability, validity, sensitivity, real time feedback, and less bias by both the experimenter and the subject [1]. Currently there is a trend to use physiological signals for many different applications and as a new form of computer interaction [2], [3]. Despite the advantages offered by this method for measuring affective reactions, there is no standardized methodology for their use in the development of affective recognition computer systems. An agreement on some conventions and guided principles would facilitate the integration of knowledge and expertise in the research community. Methodology standardization would allow comparison of results across experiments carried out in different labs, reusability of findings, and the integration of various affective recognition systems based on different approaches. We have identified five critical areas that need standardizations, conventions and guided principles, which will be reviewed in the following sections.

Area 1: Emotional Model to be Used
Many different theories about what emotion is have been proposed through the years [4]; unfortunately, there is no universal agreement on its definition or on its nature. With regards to the autonomic nervous system (ANS) activity associated to emotions, researchers still debate on how to establish a definitive model. Some theorists support that discrete emotions originate from distinct autonomic patterns [5], [6]. On the contrary, others argue that it is the perception of undifferentiated physiological arousal that originates them [7]; or that the physiological reactions are determined by the actions required from the emotional challenge [8]. Some criticisms to the physiological emotion-specificity are reported in [9], [10]. Moreover, other researchers propose an alternative dimensional model where all the sets of affective states are originated by two neurophysiological systems (one related to valence and the other to arousal) [11] [12].

Despite the non-existence of a consensus among psychologists; from a HCI perspective, only matters that affective states evoke observable physiological responses that can be identified by a computer-based system. Considering this argument only, two strategies could be considered for the development of emotion recognition systems based on physiological signals.

The first would be a pattern match approach based on proposed models about the relationship between psychology and physiology (e.g. [13]), or based on empirical findings (e.g. [14]). The dilemma is to decide which model, or empirical findings, the system should be based on. A vast number of studies document the autonomic responses to emotions; yet, some inconsistencies remain among their findings, as it can be observed on the compilation of physiological responses to specific emotions in [9].

The second option would be based on algorithms of pattern recognition and machine learning. Researchers in HCI generally follow this strategy and make use of existing, or specially created, models depending on the aim of the application or study conducted. Selecting arbitrarily a set of affective states and training a system to physiologically discriminate those specific emotions, can be seen as a practical solution for a particular problem. Nevertheless, due the nature of this strategy, the comparison of results amongst different studies is difficult as well as the integration of systems based on different approaches (see table 1).

In conclusion, the intrinsic nature of pattern matching presents favourable characteristics for standardization in HCI systems. However, its use of predefined patterns creates a controversy as there are no well-established physiological patterns for affective states so far. This has led the HCI community to privilege the use of pattern recognition and machine learning in affective recognition systems.

Area 2: Stimulus used for the identification of physiological patterns.
One reason for the inconsistency of physiological patterns among different studies might be due to the use of different types of stimulus to elucidate a particular emotion. Among the different methods in literature, one can find: staged manipulations [15], directed facial actions [16], imagery techniques [17], pictures [18], music [19], film clips [20], dyadic interaction [21], etc. When deciding about which stimuli to use for the pattern recognition stage of a HCI application, the following issues need to be taken into consideration:

I. The different types of methods to elucidate affective states offer advantages and disadvantages regarding the ecological validity and experimental control [22].

II. Due to individual differences, the same stimulus might not evoke the same emotional reactions to all participants [23]. For this reason, it is important to verify, directly with the subject, if the stimuli used succeeded to evoke the emotion intended.

III. The emotional and hence physiological reactions to a same stimulus can be different for the same person at different points in time. Multiple exposures can desensitize the subject [24].

IV. The same emotion can be experienced at different intensities depending on the context and type of the stimuli, involving different physiological reactions [25]. Therefore, the stimuli and context chosen for the recognition of patterns should be as similar as possible to the "real" stimuli and context that will be experienced later by the user.

A fundamental step towards the standardization of physiological patterns for affective states concerns the standardization of the stimulus used for their elicitation. Efforts in this direction can be found in [26], [27] and [28] where sets of pictures (IAPS), sounds (IADS) and film clips have been proposed to evoke different affective states. A comprehensive review of the methods and resources used to evoke emotions is available on the Handbook of Emotion Elicitation and Assessment [29]. In HCI, the standardization of the stimulus also needs to take in consideration the context in which the interaction will take place. Two other aspects that require well defined guidelines are: (a) the period of time for which a stimulus needs to be present to trigger a clear physiological reaction, and (b) the time needed for participants to recover from an emotional experience.

Area 3: Physiological Measures to be used

Common physiological measures used in research to identify emotions include: cardiovascular, electrodermal, muscular tension, respiration, brain activity and ocular responses. Some studies have reported the relevance of some physiological measures for particular affective states. For instance, Bradley & Lang [30] found in their experiments that facial supercilii muscle activity strongly correlates with the reports of pleasure, while skin conductance strongly covariates with the report of emotional arousal.

Regarding HCI applications, Meehan et al. [31], for example, found that changes in heart rate correlate well when evaluating stressful virtual environments. As it can also be observed in the summary of empirical work by [9] some of the findings suggest that it is possible that some physiological signals behave similarly on different emotions; however a difference in the whole set of physiological responses is appreciated. For this reason, further research is needed to determine which physiological measures provide the best results when identifying a specific emotion.

Area 4: Features to Analyze

Common features extracted from the physiological measures to identify emotions are the mean and standard deviation of the signal. Again, there are no defined guidelines. Most of the researchers in HCI based their selection of features on both previous findings and on the nature of the physiological measures selected. Heart rate, inter beat interval, amplitude, and other heart rate variability parameters, are examples of features that can be extracted from cardiovascular measures due to their underlying nature. Feature extraction methods can be used to generate new features based on transformations or combinations of the original feature. Some of the common methods include: Principal Component Analysis (PCA), Linear Discriminant Analysis, Projection Pursuit, Independent Component Analysis, Kernel PCA, PCA Network, Nonlinear auto-associative network, Multidimensional scaling and Sammon's projection and Self-Organizing Map.

It is important to note that the number of features to process will have an effect on the speed of the classifier and on the use of memory. In order to use only the features that best discriminate among the classes, methods for feature selection can be used. Examples of these include: Exhaustive Search, Branch-and-Bound Search, Best Individual Features, Sequential Forward Selection, Sequential Backward Selection, "Plus t-take away r" Selection and Sequential Forward Floating Search and Sequential Backward Floating Search. Jain et al. [32] present a review on feature extraction and selection methods. All this give the possibility of the use of many different combinations from a wide range of features. Further investigation is also needed in this area to establish which method determines the set of features that best discriminate affective states. Moreover, as the current expectations for HCI applications imply real-time responses; it is consequently highly desirable for the features to be extracted in real time.

Area 5: Models for Pattern Recognition and Classification of Emotions

The last area identified that needs standardization concerns the use of models and algorithms for pattern recognition and emotion classification. The recognition of patterns can be done using machine learning algorithms from a supervised, unsupervised or semi-supervised classification approach. In supervised learning, the training sets are already provided. In contrast, in unsupervised learning, there is no given a priori label of patterns; the system determines itself the classes based on statistical information. A combination of both labelled and unlabelled examples is carried out in semi-supervised learning [33].

There is a vast combination of possible algorithms depending on the approach used to classify the data. However, all the algorithms can be grouped on one of the three basic problems in statistical classification: (i) finding a map from features space to a set of labels; (ii) estimating the class given the training data; (iii) estimating class-conditional probabilities and then produce a class probability. A comprehensive review of common methods used in the various stages of pattern recognition systems can be found in [32]. Regarding the emotional classification in HCI, there are different algorithms that have been used such as: linear discriminant classifiers, neural networks, support vector machines, k-nearest neighbours, Bayesian networks, decision trees, etc. The main unsolved issue in this area is to determine what type of methods and algorithms provide the best results. Table 1 presents a summary of relevant studies in affective recognition for HCI, related to the five areas discussed in this paper.

Conclusion

Despite the lack of consensus among psychologists about the nature, theories, models, and specificity of physiological patterns for each emotion, psychology signals offer a great potential for the recognition of emotions in computer systems. In order to fully exploit the advantages of physiological measures, standardization needs to be established on the five key areas identified in this document. For each area, a review of the research carried out in psychology and in HCI was presented; along with the problems originated by the different methods used; and a discussion to guide further research

Table 1. *Comparison of Studies in Affective Recognition for HCI related to the five areas discussed in the paper.*

Author	Affective states	Stimulus	Physiological measures	Features	Classification	Results
[33]	No emotion, Anger, Hate, Grief, Platonic Love, Romantic Love, Joy, Reverence	Guided imagery technique	EMG, BVP, Electrodermal, Respiration	40	Combination of SFFS and FP methods	81.25% for all 8 emotions
[34]	3 Positive and 3 Negative states (with low, medium and high arousal characteristics)	IAPS	ECG, BVP, EMG, Electrodermal, Respiration, Temperature	13	Neural Net classifier	96.6% Arousal 89.9% Valence
[35]	Sadness, Anger, Fear, Surprise, Amusement, Frustration	Movie clips and math problems	Electrodermal, Temperature	Not specified	DFA, KNN, MBP	KNN: 71%, DFA: 74%, MBP: 83%
[36]	Joy, Anger, Sadness, Pleasure	Music songs chosen by participants	EMG, ECG, Electrodermal Respiration	32	KNN LDF MLP	About 80% for 3 classifiers.
[37]	Engagement, Anxiety, Boredom, Frustration, Anger.	Solving anagrams and playing videogame Pong	ECG, BVP, Electrodermal, EMG, Temperature	46	KNN RT BNT SVM	SVM: 85.81% RT: 83.50% KNN: 75.16% BNT: 74.03%.

SFFS = Sequential Floating Forward Search
FP = Fisher Projection
DFA = Discriminant Function Analysis
KNN = k-Nearest Neighbour algorithm
MBP = Marquardt Backpropagation
LDF = Linear Discriminant Function
MLP = Multilayer Perceptron Network
RT = Regression Tree
BNT = Bayesian Networks
SVM = Support Vector Machines
EMG = Electromyography
BVP = Blood Volume Pulse
ECG = Electrocardiogram

decisions. The paper ends with a table summing up the work carried out on emotion recognition for HCI, related to the five proposed areas.

Acknowledgements

We would like to thank CONACYT for their financial support and Jessica Cauchard for her encouragements and help.

References

1. Meehan, M. (2001). Physiological Reacton as an Objective Measure of Presence in Virtual Environments. PhD Thesis. Chapel Hill: University of North Carolina.
2. Nagashima, Y. (2003). Bio-Sensing systems and Bio-Feedback systems for Interactive Media Arts. Montreal, Canada. *Conference on New Interfaces for Musical Expression*: 48-53.
3. gtec. (2007). g.BCIsys - Brain-Computer Interface. Retrieved April 09, 2008, from g.tec Guger Technologies: http://www.gtec.at/products/g.BCIsys/bci.htm
4. Stearns, P. N. (1993). History of Emotions: The Issue of Change. In M. Lewis, & J. M. Haviland, *Handbook of Emotions*: 17-28.
5. James, W. (1884). What is an emotion? *Mind*, **9**:188-205.
6. Ekman, P., Levenson, R. W., & Friesen, W. (1983). Autonomic nervous system activity distinguishes among emotions. *Science*: 1208-1210
7. Mandler, G. (1975). Mind and Emotion. New York: Wiley.
8. Frijda, N. H. (1986). The Emotions. New York: Cambridge University Press.
9. Cacioppo, J. T., Klein, D. J., Berntson, G. G., & Hatfield, E. (1993). The Psychophysiology of Emotion. In M. Lewis, & J. M. Haviland, *Handbook of Emotions*: 119-142.
10. Zajonc, R. B., & McIntosh, D. N. (1992). Emotions research: Somo promising questions and some questionable promises. *Psychological Science*: 70-74.
11. Russell, J. A. (1980). A circumplex model of affect. *Journal of personality and Social Psychology*: 1161-1178.
12. Posner, J., Russell, J. A., & Peterson, B. S. (2005). The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology. *Development and Psychopathology*: 715-734.
13. Cacioppo, J. T., & Tassinary, L. G. (1990). Inferring Psychological Significance From Physiological Signals. *American Psychologist*:16-28.
14. Stemmler, D. (1989). The autonomic differentiation of emotions revisited: Convergent and discriminant validation. *Psychophysiology*: 617-632.
15. Ax, A. F. (1953). The physiological differentiation between fear and anger in humans. *Psychosomatic Medicine*: 433-442.
16. Levenson, R. W., Ekman, P., & Friesen, W. V. (1990). Voluntary facial action generates emotion-specific autonomic nervous system activity. *Psychophysiology*: 363-384.
17. Roberts, R. J., & Weerts, T. C. (1982). Cardiovascular respondign during anger and fear imagery. *Psychological Reports*: 219-230.
18. Bradley, M. M., Codispoti, M., Cuthbert, B. N., & Lang, P. (2001). Emotion and Motivation I: Defensive and Appetitive Reactions in Picture Processing. *Emotion*: 276-298.
19. Baumgartner, T., Esslen, M., & Jancke, L. (2005). From emotion perception to emotion experience: Emotions evoked by pictures and classical music. *International Journal of Psychophysiology; 34-43.*
20. Averill, J. R. (1969). Autonomic response patterns during sadness and mirth. *Psychophysiology*: 399-414.
21. Adsett, C. A., Schottstaedt, W. W., & Wolf, S. G. (1962). Changes in coronary blood flow and other hemodynamic indicators induced by stressful interviews. *Psychosomatic Medicine*: 331-336.
22. Levenson, Robert W. (2003). Autonomic Specificity and Emotion. In R.J. Davidson, K.R. Scherer and H.H. Goldsmith. *Handbook of Affective Sciences*: 212-224.
23. Cacioppo, J. T., Uchino, B. N., Crites, S. L., Snydersmith, M. A., Smith, G., Bernston, G. G., et al. (1992). The relationship between facial expressivenes and sympathetic activation in emotion: A critical review, with emphasis on modeling underlying mechanisms and individual differences. *Journal of Personality and Social Psychology*: 110-128.
24. Andreassi, J. L. (2006). Psychophysiology: Human Behavior and Physiological Response. Lawrence Erlbaum Associates.
25. Stemmler, G. (2003). Methodological Considerations in the Psychophysiological Study of Emotion. In R. J. Davidson, K. R. Scherer, & H. H. Goldsmith, *Handbook of Affective Sciences*: 225-255.
26. Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (2005). International Affective Picture System (IAPS): *Affective ratings of pictures and instruction manual*. Gainesville, Florida.: University of Florida.
27. Bradley, M. M., & Lang, P. J. (2007). The International affective Digitized Sounds (2nd Edition; IADS -2): *Affective Ratings of Sounds and Instruction Manual*. Gainesville, Florida: University of Florida.
28. Gross, J.J. & Levenson, R. W. (1995) Emotion elicitation using films.*Cognition and Emotion*: 87-108.
29. Coan, J. A., & Allen, J. J. (2007). *Handbook of Emotion Elicitation and Assessment*. Oxford university press..

30. Bradley, M., & Lang, P. (2007). The International Affective Picture System (IAPS) in the study of emotion and attention. In J. A. Coan, & A. J. B., *Handbook of emotion elicitation and assessement*. Oxford University Press.
31. Meehan, M., Insko, B., Whitton, M., & Brooks, F. P. (2002). Physiological measures of presence in stressful virtual environments. *ACM International Conference on Computer Graphics and Interactive Techniques*: 645-652.
32. Jain, A. K., Duin, R. P., & Mao, J. (2000). Statistical Pattern Recognition: A Review. IEEE *Transactions on Pattern Analayisis and Machine Intelligence*: 4-37.
33. Alpaydin, E. (2004). Introduction to Machine Learning. Cambridge, Massachusetts: The MIT Press.
34. Picard, R. W., Vyzas, E., & Healey, J. (2001). Toward Machine Emotional Intelligence: Analysis of Affective Physiological State. IEEE Transactions Pattern *Analysis and machine intelligence*: 1175-1191.
35. Haag, A., Goronzy, S., Schaich, P., & Williams, J. (2004). Emotion Recognition Using Bio-sensors: First Stepts towards an Automatic System. Lecture Notes in *Computer Science, Affective Dialogue Systems*: 36-48.
36. Nasoz, F., Alvarez, K., Lisetti, C. L., & Finkelstein, N. (2003). Emotion Recognition from Physiological Signals for Presence Technologies. *International journal of Cognition, Technology and Work*. Special Issue on Presence. 6(1).
37. Wagner, J., Kim, J., & Andre, E. (2005). From Physiological Signals to Emotions: Implementing and Comparing Selected Methods for Feature Extraction and Classficiation. IEEE *International Conference in Multimedia and Expo*: 940-943.
38. Rani, P., Liu, C., Sarkar, N. (2006). An empirical study of machine learning techniques for affect recognition in human-robot interaction. *Pattern Analysis and Applications*: 58-69.
39. Solomon, R. C. (1993). The Philosophy of Emotions. In M. Lewis and J.M. Haviland , Handbook of Emotions: 3-15.
40. Peter, C. and Herbon, A. (2006). Emotion representation and physiology assignments in digital systems. *Interacting with Computers*: 139-170.
41. Cerutti, S., Bianchi, A. and Mainardi, L. (1995). Spectral Analysis of the Heart Rate Variability Signal. Ed. Malik M. and Camm J. *Heart Rate Variability*: 63-74.

The Observer XT: a tool for the integration and synchronization of multimodal signals

Patrick H. Zimmerman[1], Liesbeth Bolhuis[2], Albert Willemsen[1], Erik S. Meyer[1], and Lucas P.J.J. Noldus[1]

[1]*Noldus Information Technology, Wageningen, The Netherlands. p.zimmerman@noldus.nl*

[2]*Adaptation Physiology Group, Wageningen Institute of Animal Sciences, Wageningen University and Research Centre, The Netherlands*

The Observer was originally developed as a professional and complete manual event recorder for the collection, management, analysis and presentation of observational data in animals [7]. However, due to its flexibility it soon became clear that The Observer was suitable for almost any study involving collection of observational data. Since then The Observer has been used in a wide range of research areas, such as applied ethology [e.g., 6], zoology [e.g., 8], entomology [e.g., 2], neuroscience [e.g., 1], psychology [e.g., 5] usability testing [e.g., 4]and sports sciences [e.g., 3].

More and more researchers combine behavioral observations with measurement of other types of data such as heart rate, blood pressure or eye movements. The benefit of combining different modalities is that you get a more complete picture of the phenomena you are studying. For instance, when you are testing the usability of your company's website, it might be useful to also use an eye tracking system. Eye fixations measured with an eye tracking system can tell you what the focus is of your participant's attention. This is very useful information that is hard to obtain from behavioral observations or questionnaires. Another example is the measurement of physiological data of a rat in an open field. The open field test is a paradigm used in neuroscience to study the effects of drugs on, for instance, anxiety or depression-like behaviors. Physiological data, e.g. heart rate or EEG data, may be of great value here. It is very well possible that on a behavioral level you do not find any effects of your drug, while your physiological data indicate that there is a significant effect (or the other way around).

The Observer XT allows you to import almost any kind of external data acquired with an external Data AcQuisition (DAQ) system, such as an eye tracking system or a physiological DAQ system to measure ECG or EEG. The only requirements for import of external data into The Observer XT are: a) the external data must be in ASCII format, b) the external data file must contain information about sample rate and c) the external data must have been sampled with a constant sample rate. It is also possible to import event data obtained with another program than The Observer, for instance uLog. Event data do not necessarily have to be observational data, but can also be R-tops from an ECG or blood pressure spikes. The only requirement is that the event data are in ASCII format and the data file consists of a column with time stamps and one or more columns with values.

Firstly, by importing external data into The Observer you achieve the integration of multimodal signals, such as observational data, eye tracking data, physiological data or video files. Secondly, you need to synchronize all signals. For example, when studying the psychophysiological response of a patient during a medical consult with a physician , you want

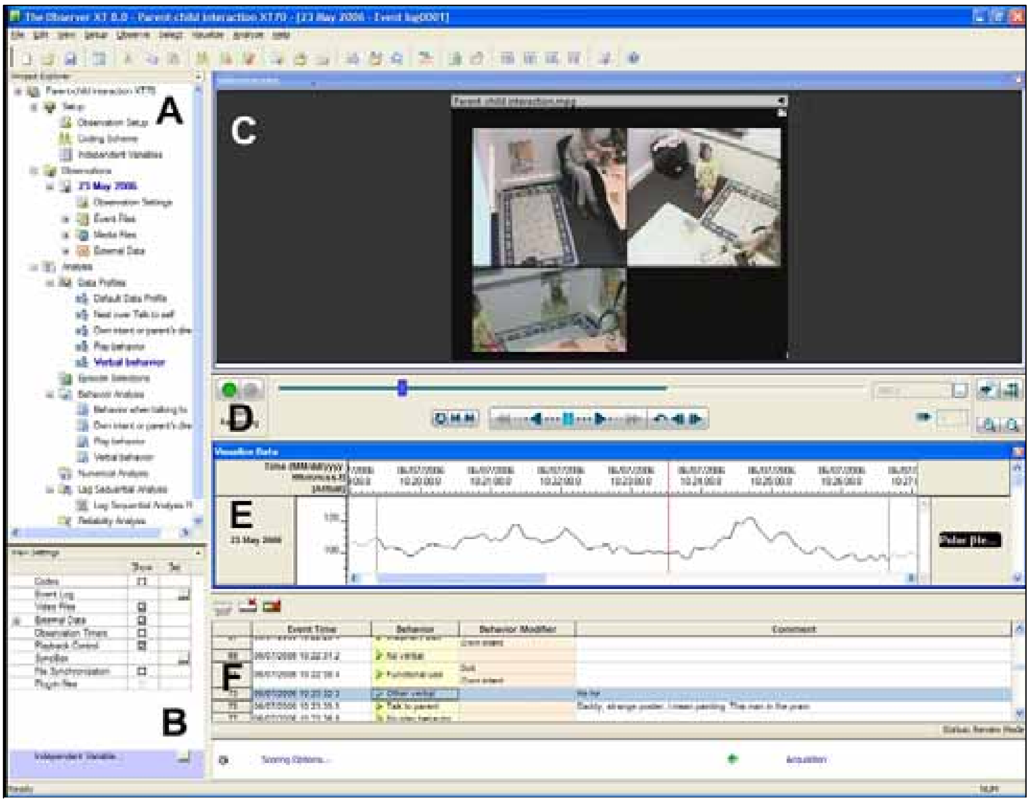


Figure 1. *Screenshot of the user interface of the Parent-child interaction sample project in The Observer XT. This sample project shows the interaction between a 6-year old girl, Suzanne, and her father. Her behavior is videotaped and her heart rate is measured using a Polar heart rate monitor. A: Project Explorer, B: View Settings, C: Video window, D: Playback Control, E: External data window with heart rate data, F: Event Log window with behavioral data.*

to record how physiological measures, such as heart rate or blood pressure, change as a result of specific types of questions asked by the physician (see the case study on our website for more details: http://www.noldus.com/site/content/files/case_studies/cs_psycho_holt.pdf). It is obvious that in this example the observational data, physiological data and video files must be synchronized, in order to see the relationship between events in the separate data sources.

The Observer XT offers several mechanisms to synchronize your behavioral data, external data and video files. You can manually set the offset for each modality to synchronize all signals. However, the easiest method to synchronize behavioral and physiological data is to carry out a live observation while The Observer XT sends out a synchronization signal to the external DAQ system on which you simultaneously acquire physiological data. The synchronization signal, with time/date information, is sampled by the DAQ system as if it is a physiological signal. Upon import of the physiological data in The Observer, the time/date information in the synchronization signal is used to automatically synchronize the physiological data with the behavioral data (see figure 1 for a screenshot from another project).

Next, you can select and visualize your data in a chart in The Observer XT; the scored behaviors and the associated physiological data are plotted against a time axis. When you play back or scroll through the observation, the behavioral and physiological data and video file(s) are played back synchronously. This allows you to visually inspect the relation between behavioral and physiological data in detail. Furthermore, you can export your behavioral data and associated physiological data to ASCII files for further analysis in other programs.

References

1. Abrahams, B.S.; Kwok, M.C.H.; Trinh, E.; Budaghzadeh, S.; Hossain, S.M.; Simpson, E.M. (2005). Pathological aggression in "fierce" mice corrected by human nuclear receptor 2E1. *Journal of Neuroscienc*e **25**, 6263-6270
2. Gols, R.; Bukovinszky, T.; Hemerik, L.; van Lenteren, J.C.; Vet, L.E.M. (2005). Reduced foraging efficiency of a parasitoid under habitat complexity: implications for population stability and species coexistence. *Journal of Animal Ecology* **74**, 1059-1068.
3. James, N.; Mellalieu, S.D.; Jones, N.M.P. (2005). The development of position-specific performance indicators in professional rugby union. *Journal of Sports Sciences* **23**, 63-72.
4. Kaikkonen, A.; Kekalainen, A.; Cankar, M.; Kallio, T.; Kankainen, A. (2005). Usability testing of mobile applications: a comparison between laboratory and field testing. *Journal of Usability Studies* **1**, 4-16.
5. Marlier, L.; Schaal, B. (2005). Human newborns prefer human milk: conspecific milk odor is attractive without postnatal exposure. *Child Development* **76**, 155-168.
6. Morrison, R.S.; Hemsworth, P.H.; Cronin, G.M.; Campbell, R.G. (2003). The social and feeding behaviour of growing pigs in deep-litter, large group housing systems. *Applied Animal Behaviour Science* **82**, 173-188.
7. Noldus, L.P.J.J. (1991). The Observer - a software system for collection and analysis of observational data. *Behavior Research Methods, Instruments & Computers* **23**, 415-429.
8. Pilz, K.M.; Quiroga, M.; Schwabl, H.; Adkins-Regan, E. (2004). European starling chicks benefit from high yolk testosterone levels during a drought year. *Hormones and Behavior* **46**, 179-192.

Connecting Brain and Behavior in Educational Research

S. R. Campbell
Faculty of Education, Simon Fraser University, Metro Vancouver, Canada, sencael@sfu.ca

Educational neuroscience is a fledgling new area of educational research that augments traditional methodologies in educational research with methods and tools from cognitive neuroscience and psychophysiology. The Educational Neuroscience Laboratory in the Faculty of Education at Simon Fraser University has been conceived and designed to help establish this new area. In this article I report on a number of projects currently underway in the lab. These projects include refining our theoretical frameworks, developing new analytical methods, and a variety of educational research applications ranging from studies in ESL anxiety and geometric image-based reasoning to metacognition and mathematics education in virtual environments. Although all of these projects are on-going or in various stages of completion, this overview demonstrates that educational neuroscience is affording new opportunities for educational research.

In this paper I briefly report on the background, conception, design, and development of the ENGRAMMETRON, a new educational neuroscience laboratory in the Faculty of Education at Simon Fraser University, a laboratory that has been established for new opportunities and directions in educational research [1]. I then report on a number of collaborative projects that are currently underway and in various states of completion. These projects include refining theoretical frameworks, developing new analytical methods, and a variety of educational research applications ranging from studies in ESL anxiety and geometric image-based reasoning to metacognition and mathematics education in virtual environments. Specifically, here, I offer brief reports on the following projects.

1) Embodied cognition
2) New methods for old problems
3) Multistable perceptions and geometric image based reasoning
4) Metacognition and motivation in self-regulated learning
5) Mathematics education in virtual environments
6) Learning biofeedback
7) ESL anxiety
8) On neuropedagogy and the importance of outreach

Embodied cognition [2] is an on-going project concerning the theoretical framework that guides and justifies all of our activities in the lab. It is a non-dualist view of cognition and learning that acknowledges the embodiment of human subjectivity. Explicating, situating, and expanding upon this view is of the utmost importance in our efforts to bring educational neuroscience into the mainstream of educational research. Embodied cognition is the basis upon which we address questions pertaining to the educational relevance of studying brain and body as part and parcel of our studies in cognition, affect, and learning. It is appropriate, then, for us to begin with a brief report on our work in this area.

Establishing a new approach to educational research, in addition to a soundly developed theoretical framework, typically involves new methodologies. This is most certainly the case for educational neuroscience. The approach we are taking to educational neuroscience in the ENGRAMMETRON involves the use of electroencephalograms (EEG), electrooculograms (EOG), electrocardiograms (EKG), eye-tracking (ET), to mention a few of the most notable data sets, integrated in a time-synchronous manner with more traditional audiovisual (AV) data sets. It is also appropriate, then, to report on some of the new methods we are working with.

One of the central topics the ENGRAMMETRON has been designed to research is the nature of mathematical cognition and learning. When a learner is looking at a geometrical diagram, that much is obvious. How do we know what part of the diagram a learner is looking at in any given moment? How do we gain insight, verbal reports aside, as to what they are thinking and when? One of the most intriguing areas of study in this regard concerns mathematical pattern recognition and mathematical concept formation. Toward this end, I report on our pilot study investigations into multistable perceptions and geometric image based reasoning.

The first major project the lab has undertaken has been a study on metacognition and motivation in self-regulated learning. Data acquisition for this project is in its final stages with over a hundred participants thus far, and data analysis is also well underway. The experimental design of this project involves participants' study of a basic theorem of number theory, and as such, the large data set collected as part of this project has implications for mathematics education research as well. I present and discuss some preliminary results from this study.

The methods we use in the ENGRAMMETRON are particularly well suited for investigating learners' interactions with visual stimuli presented on a computer monitor. With the internet becoming so ubiquitous, we are well situated for studying the latest innovations on the web. One of these is the emergence of virtual reality environments. Here I report on a project initiative to implement and study various aspects of mathematics education in such an environment.

An important dimension of educational neuroscience is to explore to what extent we are capable of placing ourselves in brain and body states and processes that are most conducive to various aspects of learning, such as memorizing, remembering, imagining, reasoning, and general states and processes associated with "brain-storming" and other kinds of problem solving activities. One avenue into such matters could be to use biofeedback and neurofeedback techniques to explore such states a processes. As a first step in this regard, I report on a project investigating how well a gaming environment can help participants learn biofeedback.

One contribution of the brain sciences over the past few years is growing evidence for the importance of affect in cognition and learning. It is now well known that affect can impede or improve learning. One of the major ways in which affect can impede learning is through anxiety. Anxieties can come in many forms, and two forms that we are most interested in understanding and unraveling are math anxiety and English as a Second Language (ESL) anxiety. In this paper, I report on one pilot we are running on ESL anxieties in Iranian women in Canada.

We are not pursuing research in educational neuroscience for the sake of pursuing research. Our aim is to make a positive difference for teachers and learners in classroom environments. There is a huge divide between the neurosciences and education, and there are great differences between neurons in the brain and kids in classrooms. Our research in educational neuroscience aims to help bridge those differences through a more informed and less speculative approach to "brain-based education" that we are referring to as neuropedagogy. Closely aligned with this approach is a viable outreach program that includes various stakeholders and interest groups concerned with neuroscience and education.

References

1. Campbell, S. R., & the ENL Group (2007). The ENGRAMMETRON: Establishing an educational neuroscience laboratory. *The SFU Educational Review* **1**, 17-29.
2. Varela, F., Thompson, E., & Rosch, E. (1991). *The embodied mind: Cognitive science and human experience.* Cambridge, MA MIT Press.

A novel method for long-term intra-cranial electroencephalographic recordings in mice - Behaviorally related brain rhythms and hypoxia-induced seizures

L Zhang[1,3], C.P. Wu1, E. Sheppy[1], M. Wais[1], M. deCampo[3], J.H. Eubanks[2,4,5],and P.L. Carlen[1,3,5]
[1]*Divisions of Fundamental Neurobiology, liangz@uhnres.utoronto.ca*
[2]*Division of Genetics and Development, Toronto Western Research Institute, University Health Network,*
[3]*Division of Neurology, Department of Medicine, University of Toronto,*
[4]*Division of Neurosurgery, Department of Surgery, University of Toronto and*
[5]*Department of Physiology, University of Toronto, Toronto, Ontario, Canada*

Introduction

Intra-cranial electroencephalographic (EEG) recordings from mice are at high demand since genetically manipulated mouse models are widely used in neurobiological research. Advanced EEG techniques have been successfully employed in mice, such as the use of micro-driver assisted multi-electrode probes [1-3] and a telemetry system [4]. However, long-term intra-cranial EEG recordings in mice remain a challenge in experimental practice. The conventional method of electrode implantation involves three to four anchoring screws being twisted through the skull around the site of implantation. Once through the skull, these screws allow electrodes to be cemented on to the animal's head via dental acrylic. Although this technique works well for adult animals when used alone or in combination with glue [5], it is difficult when employed on young mice (≤postnatal 30 days) because their skulls are not strong enough to bear the anchoring screws. Moreover, this technique requires drilling several holes through skull for placing recording electrodes and anchoring screws, and young mice or experimentally manipulated mice may be susceptible to the surgical procedure. We describe here a glue-based method for implantation of intra-cranial electrodes. Our data show that our method is simple and reliable for long-term intra-cranial EEG recordings in young as well as in aging and transgenic mice.

Materials and Methods

We constructed an electrode assembly using polyimide-insulated stainless steel wires (outside diameter of 0.25 mm). Each assembly included two recording electrodes and one reference electrode, with a total weight of 89-90 mg. The positions and lengths of recording electrodes matched the coordinates of the desired recording sites. In the present experiments, one recording electrode was placed into the hippocampal CA1 area and the other one in the contralateral parietal cortex. The reference electrode was near the parietal cortical recording site.

We used C57 black mice in the majority of the present experiments. On the day of electrode implantation, the animals were at ages of 18-25 days (young, n=23), 5-6 months (adult, n=4) or 15-18 months (aging, n=6). In some experiments, we used mice that were deficient in gene encoding methyl CpG-binding protein 2 (MeCP2, n=6). These mutant animals are considered as a model of Rett syndrome, an autism-spectrum disorder caused by the loss of MeCP2.

The animal was anaesthetized by an intra-peritoneal injection of a ketamine mixture and placed onto a stereotaxic frame for surgical operation. After skin incision, 3 small holes (0.5 mm diameter) were drilled on the skull according to the stereotaxic coordinates of recording sites, and dura underneath was opened via a fine needle. A three-electrode assembly was gently inserted into the brain and secured in place with incised skin via glue. We used a cyanoacrylate-based glue (Insta-cure+, BSI Adhesives, Atascadero, California, USA) in our experiments as it has a cure time of several seconds and high bond strength. When applied as a liquid, the glue distributed evenly in the small space between the electrode assembly and skull and thus fixed the electrode assembly onto the skull when it was cured. The surgical procedure, from skin incision to releasing the animal from the stereotaxic frame, usually takes about 15 minutes in our hands.

Animals were allowed to recover for at least 3 days before EEG recordings and other experimentations. All mice recovered fully after the surgery and showed no neurological deficits, spontaneous seizures, infections or falling off of the implanted electrode assembly while they were monitored for up to 4 months post surgery. The general behaviors of these animals were indistinguishable from those of non-implanted, naive mice.

EEG recordings were made via using extracellular amplifiers with extended head-stages (Model-300, AM Systems Inc., Carlsborg, WA, USA). The head-stage was connected to the electrode assemble via soft wires and positioned about 10 cm above the animal head. EEG signals were recorded in a wide frequency band (0.01-1,000 Hz) and amplified 1000-2000 times before digitization (digitization rates of 66 KHz, Digidata 1300, Molecular Devices, Union City, California, USA). Data acquisition, storage and analyses were done with Pclamp software (Molecular Devices).

Results and Discussion

To examine behaviorally related EEG activities, the animal was allowed to stay in its home cage and move freely during EEG recordings. When the animal moved or explored local environment, EEG recordings consistently revealed rhythmic activities in the hippocampus and low-amplitude activities in the parietal cortex. The dominant frequencies of hippocampal rhythmic activities, as determined by corresponding peaks in spectral plots, were 6.1 0.3 Hz (n=22 mice), which are in keeping with the theta rhythm previously characterized [1]. While the animal was immobile or asleep, the hippocampal EEG displayed large-amplitude irregular activities with dominant frequencies of 1.6 0.1 Hz (n=20). Large-amplitude waveforms were often noticeable in the parietal cortex during these inactive behaviors. Similar EEG activities were observed when the animal was recorded every 7-10 days for 2-3 months. MeCP2-defficient mice showed similar EEG activities, although periodic spikes (6-7 Hz) were noticeable in some of these animals. Collectively, these observations suggest that our method is suitable for stable, long-term EEG monitoring in mice. Further experiments are needed to examine the hippocampal and cortical EEG activities that are correlated with animal's performance in memory tasks such as maze or novel object recognition tests.

To examine hypoxia-induced seizures, we adopted a protocol previously used in neonatal rats [6]. An airtight plastic

Proceedings of Measuring Behavior 2008 (Maastricht, The Netherlands, August 26-29, 2008)
Eds. A.J. Spink, M.R. Ballintijn, N.D. Bogers, F. Grieco, L.W.S. Loijens, L.P.J.J. Noldus, G. Smit, and P.H. Zimmerman

chamber (Modular incubation chamber, Billups-Rothenberg, del Mar, California) was used to conduct the hypoxic episode. We made a small hole through the top part of the chamber, allowing soft wires passing through and connecting with the implanted electrodes. A rubber gasket was used to seal the hole, hence preventing air leakage. The mouse was allowed to move freely in the chamber and its EEG activities were continuously monitored before, during and after the hypoxic episode. During the hypoxic episode, the chamber was infused with 4%O2-96%N2. The duration of each hypoxic episode varied among individual animals, ranging from 3 to 15 minutes. The hypoxic episode was immediately terminated if the animal showed irregular breath, loss of posture or iso-electrical EEG signals. A similar hypoxic episode was re-applied after a 10-minute interval. The purpose of using two consecutive hypoxic episodes was to increase severity, but reduce the mortality of the hypoxic challenge. For each animal, similar hypoxic episodes were repeated 2-3 times at an interval of 7-10 days. The idea of using such repeated hypoxic challenges was to allow potential epileptogenic plasticity develop progressively in seizure-prone structures. All animals (12 young mice and 6 aging mice) showed behavioral convulsions during hypoxic episodes, including frequent shaking, head nodding, running, bell rotating and falling. EEG ictal discharges, recognized as spike waveforms that displayed amplitudes of a few minutes and duration of 10 seconds, were clearly recognizable in 10 of 18 mice examined. These data show for the first time that repeated hypoxic episodes are capable of inducing behavioral and electrographic seizures in both young and aging mice. The protocol we described here may facilitate further investigations that reveal the mechanisms of hypoxia/ischemia-induced neuronal injuries and epileptic seizures. Works are in progress in our labs to examine whether repeated hypoglycemic episodes, as a result of insulin injections, induces EEG seizures in the forebrain structures of young mice.

References

1. Buzsáki, G, Buhl, D.L., Harris, K.D., Csicsvari, J., Czéh, B., Morozov, A. (2003). Hippocampal network patterns of activity in the mouse. *Neuroscience* **116**, 201-211.
2. Jeantet Y, Cho YH (2003). Design of a twin tetrode microdrive and headstage for hippocampal single unit recordings in behaving mice. *Journal of Neuroscience Methods* **129**, 129-134.
3. Korshunov, V.A. (2006). Miniature microdrive-headstage assembly for extracellular recording of neuronal activity with high-impedance electrodes in freely moving mice. *Journal of Neuroscience Methods* **158**, 179-185.
4. Weiergräber, M., Henry, M., Hescheler, J., Smyth, N., Schneider, T. (2005) Electrocorticographic and deep intracerebral EEG recording in mice using a telemetry system. *Brain Research Protocol* **14**, 154-164.
5. Riban, V., Bouilleret, V., Pham-Lê, B.T., Fritschy, J.M., Marescaux, C., Depaulis, A. (2002) Evolution of hippocampal epileptic activity during the development of hippocampal sclerosis in a mouse model of temporal lobe epilepsy. *Neuroscience* **112**, 101-111.
6. Jensen, F.E., Applegate, C.D., Holtzman, D., Belin, T.R., Burchfiel, J.L. (1991). Epileptogenic effect of hypoxia in the immature rodent brain. *Annals of Neurology* **29**, 629-637.

Detailed automated visual tracking of biological model organisms

Ebraheem I. Fontaine[1], Alan H. Barr[2], and Joel W. Burdick[1]
[1]Department of Mechanical Engineering, California Institute of Technology, Pasadena, CA USA
{ebraheem, jwb}@robotics.caltech.edu
[2]Department of Computer Science, California Institute of Technology, Pasadena, CA USA barr@cs.caltech.edu

Efficient and rapid quantification of complex behavioral traits related to locomotion and social interactions imposes a major bottleneck on the elucidation of many interesting scientific questions. Historically, researchers have resorted to labor intensive manual digitization of relevant features to capture such data. Here, we demonstrate automated estimation of the detailed body posture of nematodes, zebrafish, and fruit flies from calibrated video. We focus our efforts on these three genetic model organisms in order to have the greatest impact on scientific discovery. These organisms benefit from an expanding array of genetic and molecular tools, making it possible to ask even more detailed questions through genotype quantification. However, techniques to automatically quantify the phenotype remain in their infancy. One reason for the lag in techniques to quantify phenotype is the significant technical challenge involved [1,7]. Our approach addresses some of these technical challenges by developing a generalized analysis framework based on geometric generative models (GGM) and a dynamic state space model (DSSM). The modular design of this model-based visual tracking algorithm allows it to function as a toolbox where the user can choose which tools are appropriate to a given experiment. This is an improvement over other tracking algorithms that are often customized to a particular organism or experimental setup. In contrast, our approach requires little re-engineering whenever the experimental parameters are changed (e.g. organism, lighting, etc.), thus facilitating rapid advancement in scientific knowledge.

An illustration of the proposed algorithm is presented in Figure 1. It consists of three major components: the GGM that defines the organism's shape and pose, the motion model that predicts the organism's motion from one frame to the next, and the observation model which updates the location and shape of the model given measurements in the image. The GGM's are continuous functions with finite parameterizations, p and q that represent the pose and shape, respectively. They are tailored offline to capture the degrees of freedom relevant to specific experimental goals. For example, in our fruit fly model, p is the spatial transformation and joint angles of the wings, while q is the body profile of the curve revolved around a center axis along with characteristic body lengths. The shape parameters are estimated offline and assumed to remain constant during tracking. Therefore, the visual tracking problem involves sequentially estimating the pose parameters, p, via measurements, z, in a discrete time dynamic state space model. Currently, we employ a Sigma Point Kalman Filter (SPKF) [9] to recursively estimate the optimal pose parameters. These SPKFs demonstrate superior performance compared to the traditional Extended Kalman Filter (EKF) when applied to nonlinear motion and observation models.

The motion models are designed to describe the displacement of the animals between frames. Our method for incorporating a-priori knowledge allows the training data to be rescaled to different velocities and consist of completely different motion patterns [10] (e.g. fruit fly wing trajectories during controlled, voluntary take off versus startle response). This a-priori knowledge of the motion restricts the search space of possible pose configurations and achieves robustness to partial occlusions. The robustness also permits the ability to track multiple organisms simultaneously [4], an essential feature when studying behavioral traits that involve interaction (e.g. mating, aggression, etc.). The output of the motion model is used as an initial condition for the algorithm to refine using measurements from the image. Given that genetic model organisms are typically filmed within controlled laboratory environments, our observation model primarily relies on edge locations along the animal's silhouette to localize our GGM.

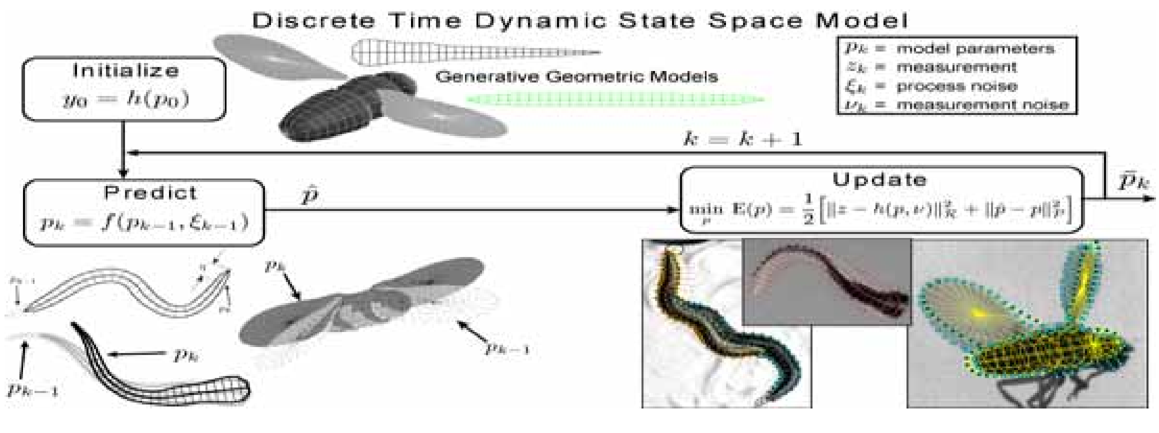


Figure 1*. Overview of Discrete Time Dynamic State Space Model (DSSM) used for automated tracking of biological model organisms. A geometric model is designed and parameterized by a finite set of parameters, p. After initializing the model to the current image sequence, the pose of the model in the current image frame is predicted based on its location in the previous frame. This predicted pose is updated using measurements from the image (e.g. edge locations). This process is repeated until the entire image sequence is tracked, yielding time varying kinematic parameters of the organism's movement.*

High contrast images permit straightforward foreground segmentation and edge feature extraction. However, when tracking multiple nematodes and zebrafish, we also incorporate a region model based on level sets [4]. Both visual cues are integrated into the framework of the SPKF.

Recently, detailed automated tracking techniques have been developed for such model organisms as *C. elegans* and *Danio rerio*. These methods achieve high levels of performance, however, they are typically limited in their scope of application. Either they are only designed for tracking a single organism in the image and cannot track multiple organisms filmed simultaneously [3,8], or the computational modeling technique is customized to a particular organism and does not extend to species of different sizes or shapes [2,11]. Figure 2 demonstrates our ability to overcome some of these limitations, with successful tracking of nematodes during mating behavior, zebrafish escape responses, and fruit flies undergoing flight initiation [4-6]. For nematodes, the mating behavior represents arguably the organism's most complex behavioral trait. Certain mutants are unable to assume particular body postures needed for successful mating, making behavioral assays with detailed posture quatification necessary to understand the underlying neuronal mechanisms of mating. Our video of zebrafish escape responses is designed to study how stiffness in the developing vertebral column influences swimming performance during ontogeny. We plan to address this question by comparing performance criteria, derived from detailed body kinematics, between wild type and mutants with stiff vertebrae. Similarly with fruit flies, to understand the underlying sensory motor control mechanisms that permit them to achieve incredibly high performance during flight maneuvers and stabilization, we must measure the detailed body and wing kinematics. Many other questions can now be asked given the wealth of information provided by these techniques. These successes illustrate the modular nature of our general approach where each module can be modified for a particular application. Current efforts are underway to harness the power of these new methods and collect large amounts of data that was previously infeasible using insufficient automated or manual techniques.

References

1. Andrews, P., Wang, H., Valente, D., Serkhane, J., Mitra, P. P., Saar, S., Tchernichovski, O., & Golani, I. (2006). Multimedia signal processing for behavioral quantification in neuroscience. *ACM International Conference on Multimedia*, 1007-1016.
2. Burgess, H. A., & Granato, M. (2007). Modulation of locomotor activity in larval zebrafish during light adaptation. *J Exp Biol* 210, 2526-2539.
3. Cronin, C. J., Mendel, J. E., Muhktar, S., Kim, Y. M., Stirbl, R. C., Bruck, J., & Sternberg, P. W. (2005). An automated system for measuring parameters of nematode sinusoidal movement. *BMC Genetics,* 6(1).
4. Fontaine, E., Barr, A. H., & Burdick, J. W. (2007). Model-based tracking of multiple worms and fish. *ICCV Workshop on Dynamical Vision*.
5. Fontaine, E., Burdick, J. W., & Barr, A. (2006). Automated Tracking of Multiple C. Elegans. *IEEE EMBS*.
6. Fontaine, E., Lentink, D., Kranenbarg, S., Müller, U., Leeuwen, J. v., Barr, A. H., & Burdick, J. W. (2007). Automated visual tracking for studying the ontogeny of zebrafish swimming. *submitted*.
7. Geneticist seeks engineer: must like flies and worms. (2007). *Nat Meth* 4(6), 463-463.
8. Geng, W., Cosman, P., Berry, C. C., Feng, Z., & Schafer, W. R. (2004). Automatic tracking, feature extraction and classification of C. elegans phenotypes. *IEEE Trans. Biomed. Eng.,* 51(10), 1811-1820.
9. Merwe, R. v. d., & Wan, E. (2003). Sigma-Point Kalman Filters for Probabilistic Inference in Dynamic State-Space Models. *Proc. of the Workshop on Advances in Machine Learning*.
10. Rosenhahn, B., Brox, T., & Seidel, H.-P. (2007). Scaled motion dynamics for markerless motion capture. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.
11. Roussel, N., Morton, C., Finger, F., & Roysam, B. (2007). A Computational Model for C. Elegans Locomotory Behavior: Application to Multiworm Tracking. *IEEE Trans. Biomed. Eng.* 54(10), 1786-1797.

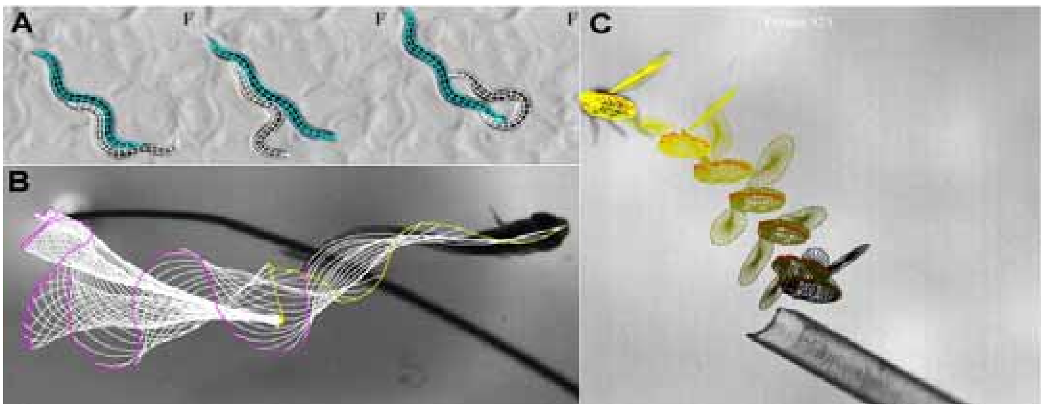


Figure 2*. Automated tracking results for genetic model organisms. The estimated location of the model is overlaid on the original camera image in order to show the fidelity of the tracking. (A) Nematodes during mating behavior, (B) zebrafish executing escape response, and (C) fruit flies during flight initiation are all possible using this approach. (A) and (B) are planar tracking from a single camera, while the 3D motion in (C) is shown from one of three calibrated camera views.*

Track3D: Visualization and flight track analysis of *Anopheles gambiae s.s.* mosquitoes

J. Spitzen[1], C.W. Spoor[2], S. Kranenbarg[2], J. Beeuwkes[1], F. Grieco[3], L.P.J.J. Noldus[3], J.L. van Leeuwen[2], and W. Takken[1]

[1] *Laboratory of Entomology, Wageningen University and Research Centre, Wageningen, The Netherlands.* jeroen.spitzen@wur.nl

[2] *Experimental Zoology Group, Wageningen University and Research Centre, Wageningen, The Netherlands*

[3] *Noldus Information Technology b.v., Wageningen, The Netherlands*

Female malaria mosquitoes locate their hosts by olfaction [1]. Little is known about nocturnal upwind navigation when the mosquito encounters host cues. This paper presents recently implemented techniques used to visualize and analyze the flight path of mosquitoes under low light conditions.

A wind tunnel equipped with a digital recording system was developed to study the flight behavior of *An. gambiae s.s.* in 3 dimensions (Figure 1). After initial tests [2], a new software package, "Track3D" [3] was developed as an add-on tool to Ethovision 3.1 to optimize automatic tracking. An air treatment system was developed and installed by Facility Services Tupola, Plant Sciences Group, WUR, The Netherlands, to ensure a constant flow of purified warm and humid 'tropical' air into the wind tunnel. Single mosquitoes were released in the wind tunnel and exposed to clean acclimatized air, human skin odours, either with or without a heat source.

The side walls and floor of the wind tunnel were constructed of black recycled polycarbonate and the ceiling was made of Lexan polycarbonate (WSV kunststoffen, Utrecht, The Netherlands). The dimensions of the flight arena are presented in Figure 1. Mosquitoes were made visible to the automatic tracking system by infrared lighting. Four infrared light units (Tracksys, Nottingham, UK) were placed at the front of the wind tunnel, facing the air flow. Each unit contained an array of 90 infrared LED's emitting light with peak output at 880 nm. To optimize lighting conditions four IR lights containing 168 LED's each (>920 nm) (Reinaert Electronics, Amsterdam, The Netherlands) were added in the same line as the Tracksys IR lights. The reflection of IR light from the mosquitoes' wings (1 wing ≈ 2.8mm) was filmed with cameras sensitive to infrared. Two Cohu 4722 monochrome CCD cameras with Fuji non-tv 9mm/f1.4 lenses were used. The cameras were attached to the ceiling of the experimental room such that their directions of view formed an angle of 40 degrees. Other camera orientations are possible, but this position minimized the reflections on the transparent top of the wind tunnel. Videos were recorded using Noldus MPEG Recorder 1.0 software (Noldus Information Technology, Wageningen, The Netherlands) that digitized the images coming from the cameras using an encoder board installed on the PC. The software saved video to a full D1 resolution, MPEG-4 video file, 704x576 pixels at 25 frames/s. The software ensured that recording started exactly at the same time for the two cameras, so the images stored in the two video files were synchronized. EthoVision 3.1 (Noldus Information Technology, Wageningen, The Netherlands) was used to analyze the video files. EthoVision 3.1 detected the mosquito on the dark background and stored the 2-D coordinates of its centre on a track file (one file for each camera).

3-D reconstruction of flight tracks

For a precise 3-D reconstruction of a mosquito flight, Track3D required input for lens correction, calibration, filtering and interpolation. Track3D accommodated the air velocity and odour plume as created for the experiments; it produced 3-D target coordinates, accuracy checks for calibration and a 3-D reconstruction. Twenty eight flight parameters were calculated and presented per mosquito track in Microsoft Excel (Microsoft Office Professional edition, 2003) output files. The target path was presented in a 3-D graph with the possibility of flight animation including different markers depending on the target position relative to the defined plume (Figure 2).

After positioning of the cameras above the wind tunnel, a calibration object of 60 x 58.5 x 57.5 cm was placed inside the flight arena. The object, made of black epoxy aluminum, had 28 markers distributed in two levels and with known 3-D coordinates. Markers on each calibration image (for each camera view) were selected in a fixed order, after which Track3D was used to determine the marker centroids. From the two sets of 2-D and known 3-D coordinates of the calibration object, Direct Linear Transformation (DLT) parameters were calculated. The calibration results were checked in two ways. First, the known 3-D coordinates were combined with the DLT parameters to calculate the expected 2-D marker coordinates. These are compared with the measured camera coordinates. Second, the 3-D coordinates of the calibration markers were calculated from the measured 2-D marker coordinates and the DLT parameters. They were compared with the known 3-D coordinates. The calibration accuracy of the used set-up was 0.5 % of the dimensions of the tracking arena.

To reduce noise the 2-D coordinates of tracked mosquitoes were smoothed by a Butterworth filter in combination with the zero phase shift routine filtfilt from Matlab 7.0 (Mathworks) with settings 'filter order 2' and cutoff frequency 8 Hz. Missing values, if not more than four at a row, were filled in by interpolated values. To this end, third order spline functions were used for all coordinates in a coherent block of data.

All flight parameter definitions are listed in the reference manual [3]. For the correct calculation in Track3D of flight parameters relative to air, the air velocity was set at 20 cm/s in the positive x direction. For each target position it was examined whether the mosquito flew inside a defined odour plume, a transition zone of 25 mm immediately outside the plume, or outside the transition zone. The odour plume was described as a cone. Its dimensions were estimated after all behavioral tests were completed by releasing smoke produced by a Safex® fog generator, F2010^{plus} (Safex-Chemie Gmbh, Schenefeld, Germany) using Safex® perfume free fog fluid. After this visualization process, the estimated position was defined by the apex, a point further on the axis, and the cone angle. Plume dimensions were estimated for treatments both with and without a heat source.

An. gambiae s.s females took off for upwind flight in 78 % of all cases tested (N=201). There was no difference in flight response between the four treatments ($P > 0.05$), however, mosquitoes exposed to the combination of odour and heat landed significantly more often at the source (47%, N=56) compared to the other treatments where landing at the source ranged between 0 and 7% ($P < 0.05$). Tracks were relatively short and in a straight line for the control treatment (e.g.

Figure 2A). The combination of human skin odours and heat elicited flights of longer duration with more crosswind behavior (e.g. Figure 2B).

Data will be presented to demonstrate the relevance of the system as a tool for the study of host orientation in mosquitoes.

Discussion and Conclusion

The automated tracking system successfully recorded and analysed flight behavior of nocturnal mosquitoes whose principal orientation cue for host recognition is smell [1]. The technical constraints associated with the recording of a fast flying small object, moving in 3 dimensions, were met resulting in a tool with which the behavioral parameters of mosquitoes and other insects can be accurately recorded and analysed. Track 3D is commercially available from Noldus Information Technology (www.noldus.com).

Acknowledgements

The authors like to thank all the people of Facility Services Tupola Plant for their input during the development of the wind tunnel and air treatment system. Leo Koopman, Frans van Aggelen and André Gidding are acknowledged for rearing the mosquitoes.

References

1. Takken, W. and Knols, B.G.J. (1999). Odor-mediated behavior of Afrotropical malaria mosquitoes. *Annual Review of Entomology*, **44**, 131-157.
2. Beeuwkes, J., Spitzen, J., Spoor, C.W., van Leeuwen, J.L., Takken, W. (2008). 3-D flight behaviour of the malaria mosquito *Anopheles gambiae* s.s. inside an odour plume. *Proceedings of The Netherlands Entomological Society Meeting,* **19**, 137-146.
3. Noldus Information Technology. (2007). *Track3D.* Software for visualization and analysis of animal movement in 3D – Add-on module to EthoVision R 3.1. *Reference Manual Noldus Information Technology b.v.,* 68 p.

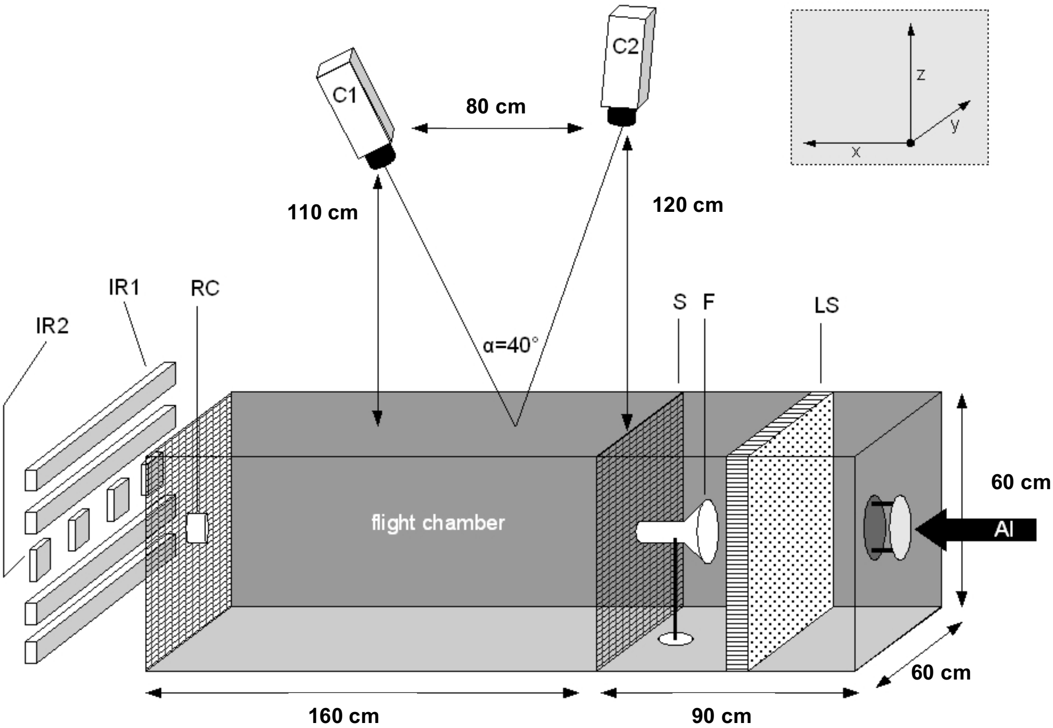


Figure 1. *Wind tunnel set-up. Air inlet (AI), lamination screen (LS), glass funnel containing heat element (F), mesh screen (S), release cup (RC), cameras (C1,2), IR lights type 1 (IR1), IR lights type 2(IR2). Mosquitoes were in view of both cameras up to 60 cm downwind from the mesh screen (S).*

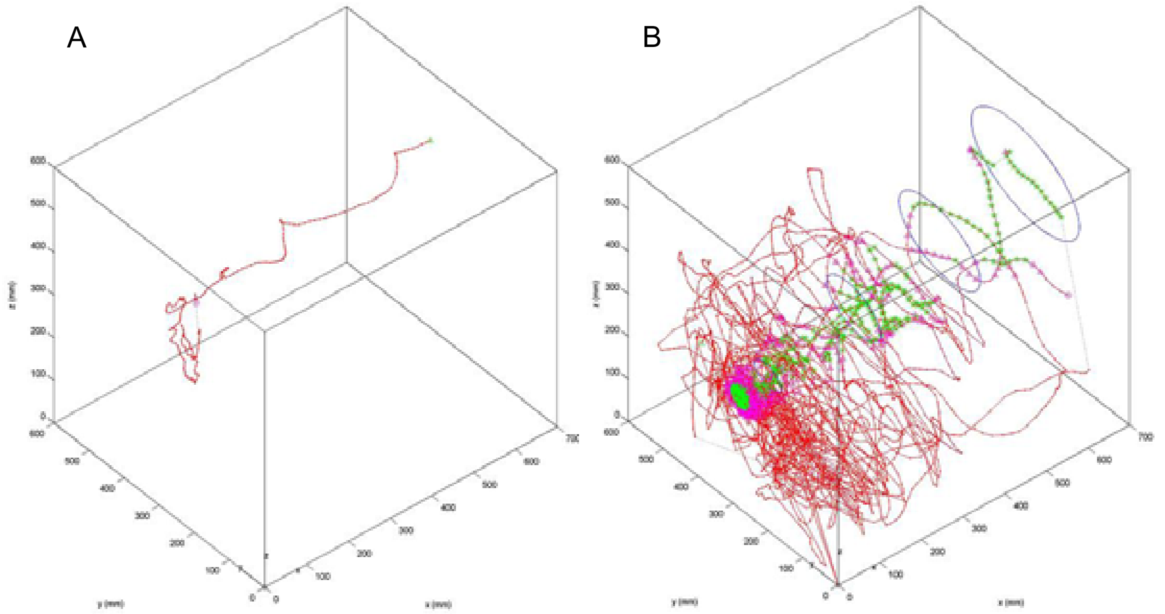


Figure 2*. 2A shows a flight track of An. gambiae s.s. exposed to clean acclimatized air only. 2B shows a track of a mosquito flying upwind towards human skin odour in combination with a heat source. (Red) dots indicate samples outside the cone and buffer zone, (magenta) triangles are used for samples within the buffer zone and (green) stars indicate that the insect is tracked within the defined odour plume.*

Real-time tracking of multiple targets using multiple laser scanners

Adam Feldman[1], Summer Adams[1], Maria Hybinette[2], and Tucker Balch[1]
[1]*College of Computing, Georgia Institute of Technology, Atlanta, USA, storm@cc.gatech.edu*
[2]*Computer Science Department, University of Georgia, Atlanta, USA*

Tracking humans, robots and animals is becoming increasingly important to analyze and understand behavior in domains ranging from biology to computer vision and robotics research. We propose a new and reliable mechanism that simultaneously and automatically tracks the locations and the number of multiple animals, objects or people (hereafter, 'targets') in a dynamic environment, indoors or outdoors, in uncertain lighting conditions as they move rapidly through the environment over time. We use multiple laser range finders (or ladars) to overcome deficiencies of computer vision [1,2] such as dealing with difficult lighting conditions, potentially heavy computational load of frame-range image processing and distinguishing foreground from background. In contrast to computer vision, ladars are more reliable because they are less susceptible to 'false positives' and 'false negatives', yet provide very high spatial accuracy. Ladars have been used in other areas of research (e.g., [3]-[5]) but they often do not address tracking of multiple, fast moving, interacting targets, and most existing research relies on a single ladar sensor.

We use multiple ladar at different viewpoints to address occlusion and cover a large area at a greater density. Although a single laser placed at a location provides only one view, and is therefore unable to see objects hidden from that view, multiple lasers overcome this problem by providing additional viewpoints from which to see these 'hidden' objects.

Before tracking begins a global 'picture' is generated from the raw data collected from the multiple sensors through temporal and spatial registration. Temporal registration is accomplished by matching data from each ladar with the most similar timestamps. Spatial registration occurs by transforming each ladar's point of view into a global frame, superimposing each set of data onto one another. After the data is registered, the background must be subtracted. The background consists of all laser hits that are not caused by any of the targets being tracked. It is made up mostly of stationary objects (e.g., the wall, chairs, etc.). Subtraction occurs by eliminating data points at and beyond (in polar coordinates relative to each ladar) the location at which the majority of data points occur, across all the frames.

Tracking resolves the location of a single target over time. This is challenging because of both (partial) occlusion and imperfect background subtraction, which result in difficulties identifying the 'interesting targets' in a given frame. Further, the data association problem, the ability to correctly associate a given target with itself over time, is especially difficult when multiple targets are in proximity to each other. To solve these difficulties, a tracker must first determine which clusters of laser hits (i.e., the grouping of laser hits which represent a target) in a given frame correspond to one of the interesting targets. Second, it must recognize these clusters from frame to frame to build tracks representing that target over time. Our tracker accomplishes this in parallel, using information about clusters found in one frame to find the corresponding cluster in the next.

First, the tracker must identify all clusters in each frame. To begin processing a new frame, each current cluster center (from the previous frame) is considered before new centers are created. A current cluster center is updated as follows. A circle is placed around the cluster center. The radius of this circle (the 'cluster radius') is designed to approximate the size of the target being tracked, so as to best enclose all laser hits related to that target, while excluding all other laser hits.

All encircled laser hits now make up the cluster being updated. The cluster center is then moved to the geometric center of the updated cluster. A circle is drawn around this new cluster center, and the list of laser hits is updated to only include those that are inside the new circle. This cycle of updating the included data list and moving the cluster center is repeated until there are no changes, resulting in the fully updated cluster being added to the list of clusters for this frame, and the included points being removed for further consideration. Figure 1 details this process, which is applied to all clusters in each frame. If a cluster is found to contain no data points in the new frame, it will 'die,' not being propagated into any future frames.

Once all previous cluster centers have been updated (if there

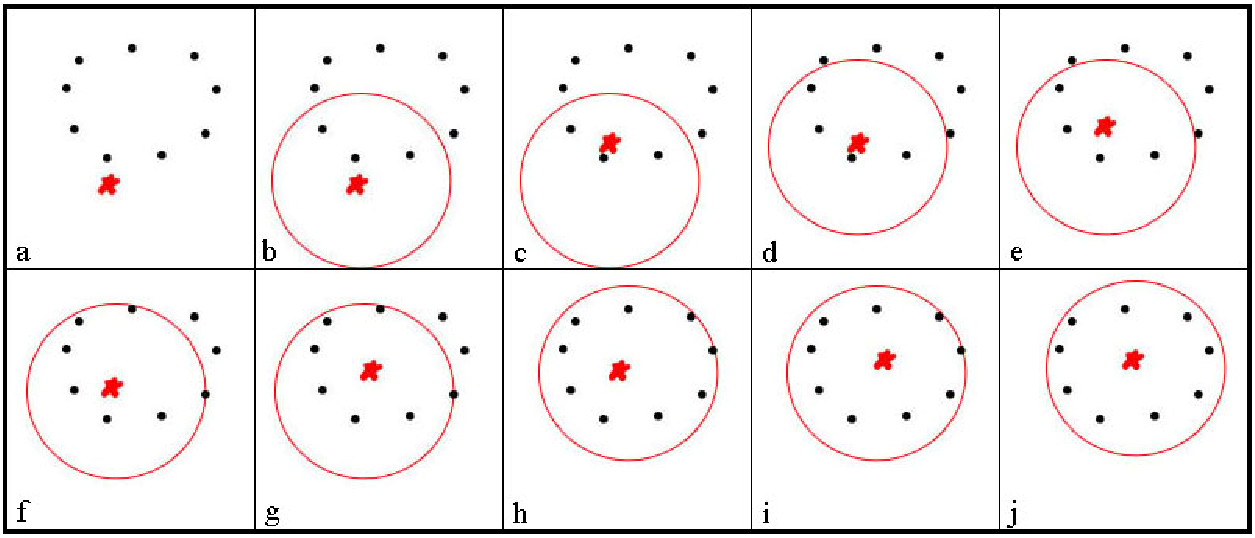


Figure 1. *Part (a) starts with the current data and the previous frame's cluster center. In (b) we draw the cluster, and then in (c) we move the cluster center to the geometric average of the three included points. Parts (d) through (j) show this process iterating until no change of points occur and we have the final cluster for this frame.*

were any), individual remaining laser hits are considered. One laser hit is chosen arbitrarily, and a cluster center is placed at that laser hit. The above process is repeated with the newly created cluster center. This procedure of creating and updating cluster centers is reiterated until no laser hits remain in this frame. Figure 2 shows the final results of a sample frame.

An efficiency of this system is that it builds tracks automatically as it identifies the clusters present in each frame by using the previous frame's cluster centers as starting locations. A cluster center is connected to other cluster centers across time to form a track only if it is in fact the same cluster center which has moved from one frame to the next, as described above. To help ensure that tracks do not mistakenly represent multiple targets, the tracker checks for situations in which multiple clusters are close to each other. In such cases, both tracks will be ended and the clusters will be assigned new tracks.

Sometimes clusters will be made up of background or noise, not actually representing targets. These spurious clusters are eliminated by the tracker refusing to output any cluster centers that are represented by fewer than a minimum number of laser hits. Note that the upper left cluster in Figure 2 was correctly eliminated in this phase.

Figure 2. *Results of processing this frame. Dots represent laser data. Circles represent generated clusters.*

References

1. S. Sclaroff and R. Rosales (1998). Improved tracking of multiple humans with trajectory prediction and occlusion modeling. *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition, Workshop on the interpretation of Visual Motion* (Santa Barbara, USA, 1998).
2. J. Perš and S. Kovačič (2000). Computer vision system for tracking players in sports games. *Proc. of the First Int'l Workshop on Image and Signal Processing and Analysis (IWISPA 2000, Pula, Croatia)*, pp. 81-86.
3. E. Prassler, J. Scholz, and A. Elfes (1999). Tracking people in a railway station during rush-hour. *Proc. of the Computer Vision Systems, First International Conference (ICVS-1999, Las Palmas, Spain)*, pp. 162-179.
4. A. Fod, A. Howard, and M. J. Mataric (2002). A laser-based people tracker. *Proc. of the IEEE International Conference on Robotics and Automation (ICRA-2002, Washington, DC, USA)*.
5. A. Panangadan, M. J. Mataric, and G. S. Sukhatme (2004). Detecting anomalous human interactions using laser range-finders. *Proc.of 2004 IEEE/RSJInternational Conference on Intelligent Robots and Systems (IROS-2004)*, pp. 2136-2141.

Imprinting with morpholine and use of silicon base tags on behavioral tracking in fish maze

Homayoun Hosseinzadeh Sahafi
Iranian Fisheries Research Organization, Po Box 14155-6116, No.297, Fatemi Ave., Tehran, Iran,
h_ hosseinzadeh@yahoo.com

Abstract

There are several methods for studding behavior of fish fry's and fingerlings in in the natural and artificial conditions Many types of maze were described for taxis, behavioral studies and imprinting as well as odor responses in anadromous fishes[1,4,5].

Commercial fishes in south Caspian Sea, such as sturgeon fishes and Kutom (*Rutilus frisi kutom*) has been exploited from 1970 up to now. More than 200 million fry and fingerlings were released into the Caspian sea in 2006. Although huge efforts were taken by fisheries sector for restocking the resources, but return rate still is unknown. There are many problems for return rate and homing of Kuum, such as pollutions, reconstruction of rivers, etc .that may be caused by mismatch entrance of fishes affected with unknown odors. Imprinting of fishes with chemical enhancers such as morpholine and phenylethyl alcohol were described by Sholz et al. . 1975, 1978[2,3]

Hexagonal maze were used for studding the imprinting behavior of Kutom fingerlings (No. = 180) during April to August 2006 (see Figure 1.). Samples were taken from fingerling cams out from natural Broodstoks in Khoshkrud river (North of Iran) that spawned on the river water .Each trials were covered 30 specimens in tree weights (1-1.5 g. , 2-2.5g. , 3.3.5 g.),separated by fluorescents tree color flexible silicon based tags(ref.44372) compare to tree color visible external plastic tags .Water flow (natural river water) supplied by peristaltic pump to the branched chambers(L. =100cm.) running downward to the center of apparatus and collected in the middle of the main chamber (r. = 30cm)[5]. Samples were exposed to different doses of morpholine (0 PPM, 5 x 10^{-5} PPM, 7 x 10^{-5} PPM) in hatching time for 30 days. Morpholine (2x 10^{-5}) were emited to the end chamber of each interval branches. Experiments were down in tree replicates.

Results show that there was significant difference between three weights in response to morpholine(see Figure 2.). Also the entrance rate of fingerlings to the emited and non emited race ways were significantly differenced ($P < 0.001$). Fingerlings witch exposed to 7 x 10^{-5} PPT in hatching time were fined to grate attraction (2> in each end chamber) to water flows with morpholine compare to the other concentrations ($P < 0.001$). Fishes with silicon tags were determined and traced by UV light during the experiments. This implanted flexible tag allows fish to have unrestricted swimming compare to external plastic tags. It is suggested to use implanted flexible tags for fine experiments with small fishes in the case of migratory and homing behavior studies. Also it would be suggested that using of morpholine as an imprinting factor would be effective in homing behavior of kutom in artificial experimental condition.

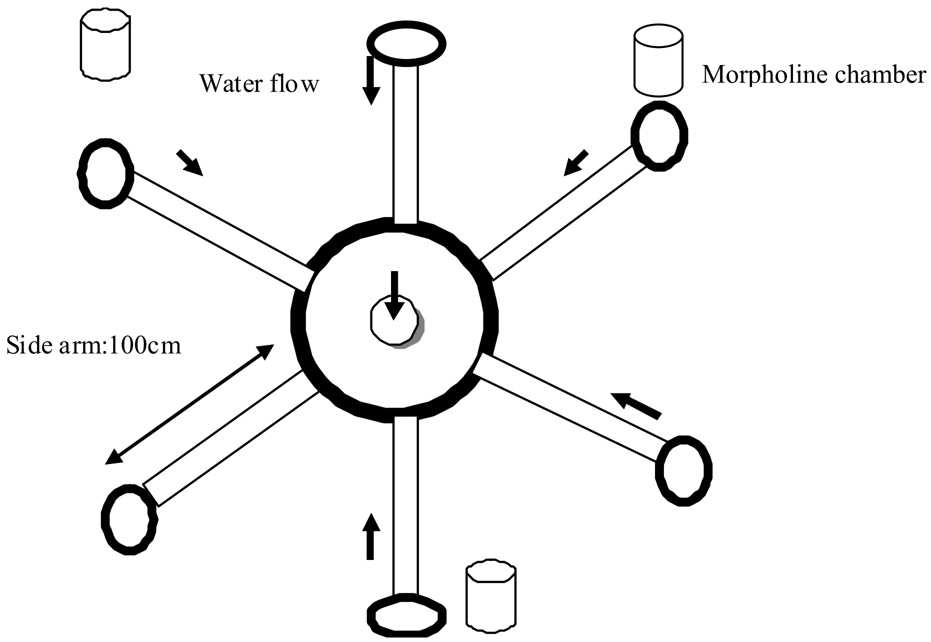


Figure 1. *Schematic design of the hexagonal maze for morpholine test*

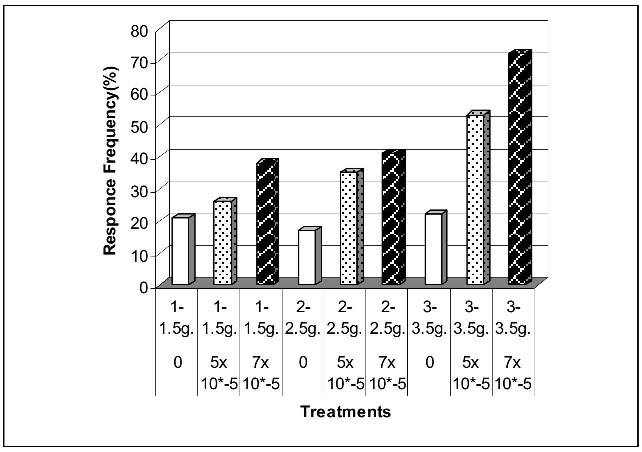


Figure 2. *Behavioral response of Kutum Fingerlings to morpholine in different weights*

Key words

Morpholine,Imprinting, Rutilus frisi kutom, Silicon, Tags

References:

1. Ricker, W.E.(1972). Hereditary and environmental factors effecting certain salmonid populations. p.27- 160 in: R.C. Simon and P.A. Larkin (editors) *The stock concept in Pacific salmon. In H.R. Macmillan Lectures in Fisheries*. University of British Columbia, Vancouver, British Columbia.
2. Scholz, A.T.,. Gosse C.K,. cooper J.C,. Horrall R.M,. Hasler A.D. Poff , R J,. Daly. R.I (1978).Homing of rainbow trout transplanted in Lake Michigan: a comparison of three procedures used for imprinting and stocking. *Transactions of the American Fisheries Society* **107**:439-444.
3. Scholz, A.T.,. Horrall R.M, Cooper J.C.,. Hasler A.D,. Madison D.M,. Poff R.J Daly R.I (1975). Artificial imprinting of salmon and trout in Lake Michigan. Wisconsin Department of Natural Resources Fisheries Management Report No. 80: 46 pp.
4. Dittman H A.,Quinn TP., Nevitt GA.,(1996),Timing of imprinting to natural and artificial odors by coho salmon*(Oncorhynchus kisutch)* , *Can. J. Aquat. Sci*. **53**: 434-442
5. Hughes RN.,Blight CM., (1999),Algorithmic behaviour and spatial memory are used by two intertidal fish species to solve the radial maze, *Animal Behaviour* **58**, 601-613

Active Shape Models for Mouse Behavior Classification

S.R. Campbell

Kuzer Co., Miami FL USA, kuzershan@yahoo.com

Measuring the behavior of mice is of prime interest in much genetic, toxicological, pharmacological and cognitive research (see for example [2,3]). Effective automated means of performing this task exists for several types of behavior in somewhat restricted environments. To perform accurate detection and measurement of such behaviors as grooming, rearing, stretching, and digging, in the more common environment of the laboratory cage with bedding material, more sophisticated means of extracting the mouse contour from the image are necessary.

In this project, we use active shape models (ASMs) [1] to segment the mouse in each frame of video data. ASMs are based on a collection of points used to represent the contour of the mouse. A number of images are labeled by hand to serve as training data to create the model. Statistics are extracted based on the distances between all the points. A mean shape is created and deviations from that mean are represented by the eigenmodes extracted from the training data. For example, the largest eigenmode of shape deformations for the mouse contour is that of left and right head turns. During ideal operation, the ASM is positioned near the object, at each point along the ASM, edge detection is performed and the points along the contour are moved towards the strongest edge locations. The motion of the points is constrained by the eigenmodes and coefficients of the shape model, i.e. the coefficients of the ASM are adjusted slowly, and constrained in their magnitude by the already computed eigenvalues. In this fashion, the ASM does not arbitrarily fit any edges found in the image, but those that best represent typical mouse shapes.

Given that ASMs implicitly label each point on the contour, the model has the capability to easily represent ears, nose, eyes, and trunk of the mouse. This detailed information enables us to acquire object features such as the face orientation relative to that of the body, as well as estimate the area of the head relative to that of the body, a feature is useful in determining when the mouse is rearing.

The ASM by itself is not powerful enough to perform segmentation and tracking of mice in video images. Several additional algorithms are necessary to aid the ASM. For example, at each iteration of the ASM, the points of the ASM move towards the strongest edge location in a direction perpendicular to the contour. However, this direction changes as the contour moves, and as the mouse edges move from frame to frame. It is possible that the edges become bunched up, which can lead to other problems, such as the formation of loops – a topological defect that while physically impossible, is not inherently prevented by the ASM. To prevent the majority of these topological problems, we reposition each point on the ASM so that the distance between them is uniform.

Also, the spatial search for edges performed by the ASM is limited, so when the mouse moves quickly, the ASM can “fall off” the mouse. To prevent this from occurring, the ASM is initially repositioned on each frame based on motion analysis. Thus the head and rear of the mouse are moved towards motion typically exhibited by the head and rear of the mouse. The color of the mouse and of the background must be known to appropriately distinguish forward motion of the head from forward motion of the rear. Retropulsion is a potential source of confusion, but this motion is uncommon, and all cases thus far encountered have been adequately resolved by the spatial search performed by the ASM.

The ASM suffers from another problem associated with this particular application. The mouse is highly deformable, and the precise location of the ASM is dependent both on the information in the current frame, as well as the location of the ASM in the previous frame. The ASM can shift or rotate around within the contour of the mouse. To aid the ASM, we use the ears of the mouse, which are distinctively shaded from the fur on the mouse body, to more correctly position of the ASM within the mouse contour. We created an ad-hoc algorithm specifically to find the ears of the mouse, starting with spot detectors, and eliminating the many false positives that arise by examining all pairs of spots and requiring them to have a specific distance between them, as well as a specific and fairly uniform brightness between them. Given the highly textured bedding material, we can accurately find the ears of black mice. An example of this is shown in Figure 1 (left) and also for white mice (middle). A similar algorithm was used to find the eyes of the white mice. Brown mice tended to have

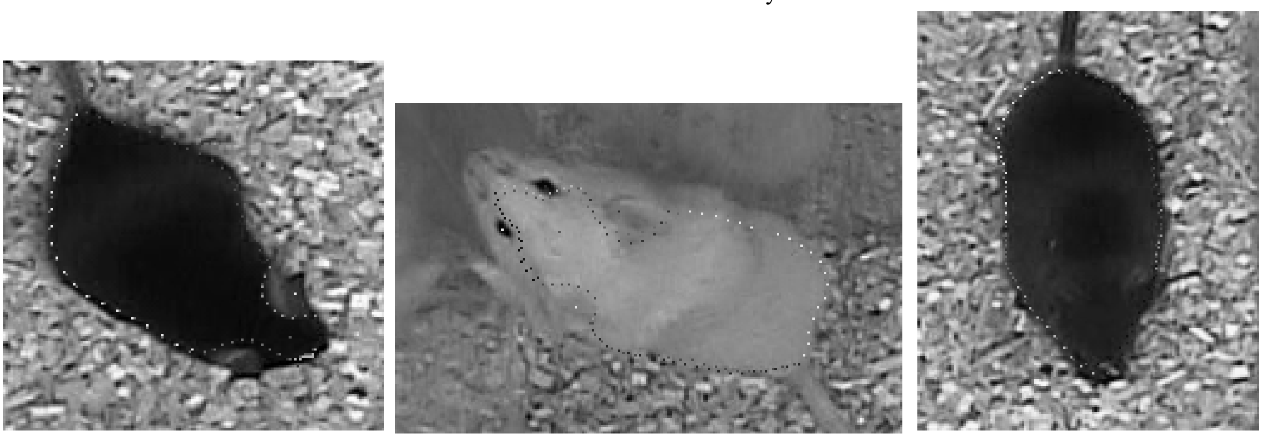


Figure 1. *Images of black, white, and brown mice (left to right) and the position of the active shape model. Differently shaded dots along the contour represent different portions of the body, left trunk, left ear, left face, right face, right ear, and right trunk. The locations of the ears are denoted by a single dot. In the brown mouse, the ears are not detected due to their exceptionally low contrast and the ASM has not accurately positioned the left and right ear contour portions.*

ear colors too similar to the rest of the body for this approach.

As can be inferred, the image information for brown mice was not as accurate as that from white and black mice. Accordingly, the features extracted from the contour were no as accurate (as can be seen in Figure 1 (right) nor were the resulting behavioral classifications.

Using the modified ASM described above, we were able to accurately segment single mice in cages and on pine chip bedding material for approximately 100,000 frames of video data. From this data we extracted a number of features, such as velocity, area of face, area of body, max thickness of trunk, etc. Using a linear discriminate analysis (LDA) (see 3), we were able to classify mouse behavior into 5 different classes, inactivity, activity, grooming, partial body motions, and rearing.

To understand the quality of the LDA classification, we compare the LDA results with those of three human raters for ~16000 frames of data. We used Williams' index 5 to compare these categorical ratings and find that our algorithm yields excellent agreement (I_n = 1.02) with those of the raters. In fact, it appears that one human rater in particular differed from the group ratings by a significant margin (I_n = 0.76); we conjecture that he may have misunderstood some portion of the instructions.

In summary, we find that unaided, the ASM was not adequate for tracking and segmenting mice in video. With assistance from motion information, as well as from the distinctive features of the mouse ears and eyes, a modified ASM could accurately segment mice for the approximately 100,000 frames of video used in our project. The data obtained from imagery was of high enough quality such that a linear discriminant function was able to categorize the mouse behavior at a high level of agreement with three human raters.

References

1. Cootes TF, Cooper D, Taylor CJ, and Graham J, (1995) "Active Shape Models - Their Training and Application." *Computer Vision and Image Understanding*, **61**(1), 38-59.
2. Crawley JN (2000) *What's wrong with my Mouse? Behavioral phenotyping of transgenic and knockout mice*, John Wiley & Sons, Inc. Publication, New York.
3. Hastie T, Tibshirani R, and Friedman J (2001), The Elements of Statistical Learning: Data Mining, Inference and Prediction, Springer-Verlag, New York NY, USA
4. Nolan PM Peters J Strivens M Rogers D *et al* (2000), A systematic genome-wide phenotype-driven mutagenesis programme for gene function studies in the mouse, *Nat. Genet.* **25** 440-443
5. Williams GW (1976), Comparing the Joint Agreement of Several Raters with Another Rater, *Biometrics* **32**, 619-627.

Real time measurement of social networking behavior using wireless sensor mesh networking technology

Richard L. Amado

C.H.E.S.S. BV Haarlem, The Netherlands, rick.amado@happytag.com

Current applications for wireless tracking and tracing usually focus on localization of absolute physical position and require a fixed infrastructure. We will present a proposal for using wireless sensor mesh networking technology to enable measurement of relative social location, clustering density, relative movement and social interaction. The technology requires little or no fixed infrastructure and can be applied independent of scale. We will describe the capabilities and constraints of the technology, and will discuss applications envisioned for marketing, live events and retail.

Problem scope: Low cost scale-free measurement of social clustering behavior in real time

Technology: Wireless sensor technology, ad hoc mesh networking using a gossip protocol

The technology we will present is based on 2.4Ghz wireless sensor hardware. The sensors form an ad hoc mesh network when in range of each other – approximately 30 meters. The network protocol is based on Gossip, a broadcast-only protocol. The combination of wireless sensors with Gossip creates two interesting inherent properties:

- The technology is scale free, ambient and completely ad hoc, allowing unlimited, unknown and varying numbers of sensors in the mesh.
- The sensors keep track of their own neighbors without central control

This self-monitoring wireless mesh lends itself to measuring social clustering behavior, especially larger gatherings – festivals, stadiums, amusement parks and congresses. The technology can be applied in an unobtrusive and anonymous way to track social behavior of individuals in a crowd.

In a congress setting, the technology can be implemented into the badge. The result is the ability to measure, in real time, social traffic and clustering.

With the addition of a near field technology in the sensors, and a method for wearer input, the system can facilitate interaction and exchange of information based on proximity.

We have developed a commercial platform – "HappyTag" which puts these properties into action for live events and retail. HappyTag facilitates interaction, transaction and is context aware.

Figure 1. *Sensor node*

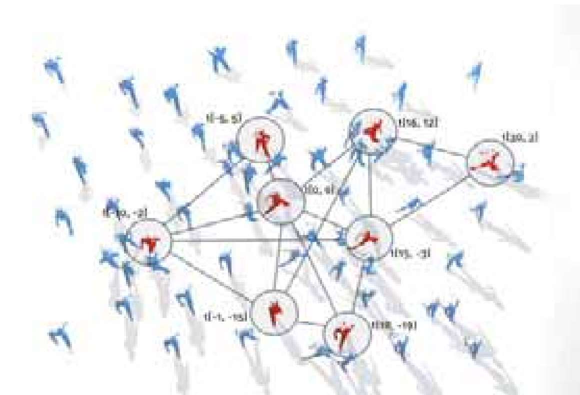


Figure 2. *Mesh of sensors in social context*

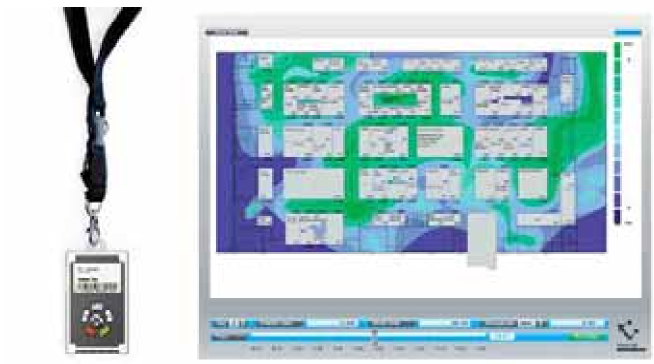


Figure 3. *Congress Badge containing sensor. View of traffic density inside a congress venue*

Figure 4*. HappyTag platform hardware components*

Automated scoring of novel object recognition in rats

K. Rutten[1], O. Reneerkens[1], H. Hamers[2], A. Şık[2], I.S. McGregor[3], J. Prickaerts[1], and A. Blokland[1,2]
[1]*Faculty of Health Medicine and Life Sciences, Brain and Behaviour Institute, Maastricht University, The Netherlands*
[2]*Faculty of Psychology and Neuroscience, Maastricht University, The Netherlands, h.hamers@psychology.unimaas.nl*
[3]*School of Psychology, University of Sydney, Australia*

Introduction

The object recognition test (ORT) is a behavioral test that is widely used to examine animal memory performance. This test was first described by Ennaceur et al. (1988) and has been used in many different variations ever since. A major drawback of the ORT is that the scoring is done by manual scoring, which can be liable to subjectivity. Various attempts have been made to automate the scoring of the rats, but to our knowledge the use of specific software has not been published yet. Recently, we developed an automatic scoring program that reliably tracks the nose of the rats. The aim of the present study was to compare and validate this scoring program by comparing the automatically obtained scores with the manually obtained scores of two experienced independent observers.

Methods

Three-month-old male Wistar rats were used. They were tested during the dark phase. The novel object recognition test was performed as described previously (Prickaerts et al. 1997). The apparatus consisted of a circular arena, 83 cm in diameter. Half of the 40 cm high wall was made of grey (RAL 7035) polyvinyl chloride, the other half of transparent polyvinyl chloride. A set of either big or small objects, each consisting of 4 different shapes, was used for the testing of the animals. Exploration to an object was defined as follows: directing the nose to the object at a distance of no more than 2 cm and/or touching the object with the nose.

Automatic detection

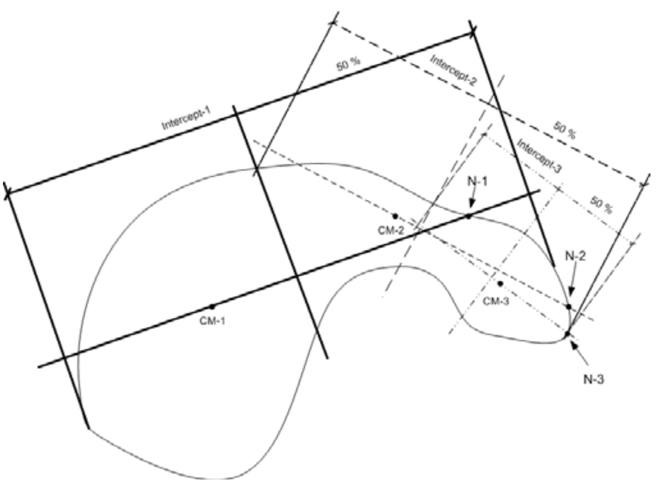


Figure. 1.

The detection algorithm of the nose was as follows. The maximal intercept of the rat was determined. Subsequently the Center of Mass (CM-1) was determined and the line of the maximal intercept was shifted so that the maximal intercept line crossed the Center of Mass. This resulted in three different X-Y coordinates: Center of Mass, intercept-body crossing front, intercept-body crossing back. Then, a perpendicular was placed on 50% of the intercept line. The perpendicular divided the body in two parts and the Center of Mass was again calculated for the front part (CM-2). The point where the maximal intercept crossed the perimeter of the body was regarded as the position of the nose (N-1, see figure.1). Using this algorithm, which was executed in a cycle of 40 ms, the nose gradually shifts to the actual nose position (see shifts from N-1 (solid lines) to N-3 (dotted-dashed lines) in figure. 1). This process is dynamic since the X-Y coordinates of the previous calculation are compared with the new input of X-Y coordinates. Thereby, this process continuously leads to the detection of the actual nose position.

Data

Table I. *Inter-observer correlations for the measure e1, e2 and d2. (e1= exploration time in trial 1; e2 = exploration time in trial 2, d2= discrimination index).*

Conditions	Objects	r^2 e1	r^2 e2	r^2 d2
SAL	small	0.95	0.95	0.72
	large	0.90	0.75	0.78
SCOP	small	0.84	0.93	0.85
	large	0.77	0.80	0.61

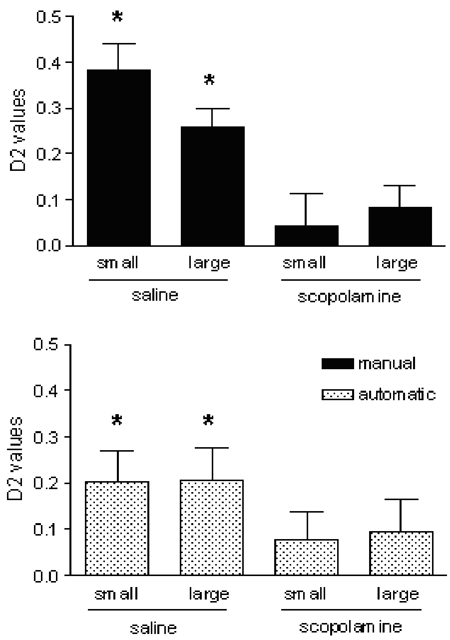


***Figure 2.**. Memory performance of rats after manual scoring (top panel) and automatic scoring (lower panel).*

Conclusion

Object recognition performance of rats can be done via automatic scoring

3D Facial model generation for surgical planning with 3D scanned data and conformal mapping

Ahmed Mostayed[1], S. Kim[1], Mynuddin Mazumder[1], and Se Jin Park[2]

[1]*Department of Electrical Engineering, Kongju National University, South Korea, shaibal125@ yahoo.com, skim@kongju.ac.kr, mynudding@ yahoo.com*

[2]*Korean Research Institute of Standards and Science (KRISS), South Korea, sjpark@kriss.re.kr*

Abstract

One of the most important issues in facial surgery modelling is to access the quality of the modelling of the face after surgical operation. This paper presents a method for modelling of face based on 3D head scan data of the patient acquired before facial surgery. The technique is made highly efficient and scalable by partitioning the 3D head model into distinctive facial regions and using conformal mapping to reshape it. At first facial features are extracted from 3D head data. A generic mesh is individualized by correspondence matching from those feature vertices using conformal mapping.

Introduction

In recent years, modern people specially the young generation has been seen to be very eager to look good and attractive and in near future they may will be wanting to be look alike some of their idols or heroes. As a result customized face plastic surgery has received increased interest among doctors and the people concerned. The adaptation of a head model to scan data is performed in two dimensions by [1, 2]. The approach taken in [3] is based on variational modelling, operating directly on a B-spline surface in three-dimensional space. The three-dimensional (3D) human face creation and modelling are important because of their wide applications in maxillofacial surgery, such as computer-aided simulation surgery [4], the 3D computer-assisted custom-designed implants [5] and video imaging prediction before orthogenetic surgery [6]. Traditionally, the face plastic surgery design methods are drawn on bare eyes and skilful hands. But we suggest it is essential to take 3D measurement of head at the first step for customizing face surgery and there fore utilize cross section head data [7].

This paper proposes a novel procedure to design the face structure for plastic surgery consisting of the following: (a) three dimensional data from a 3D head scanner; (b) an implementation of the Schwarz-Christoffel algorithm in Matlab®; and (c) a conformal mapping algorithm. First one has to take the three dimensional head data by means of a 3d head scanner. Then several cross-sections are taken from different regions of the 3D data. Next conformal mapping is performed on each cross section with Schwarz-Christoffel toolbox.

3D Head measurement system and data

The 3D head data is the three dimensional coordinate points on the head surface. The raw data obtained from the head scanner is a 3-by-3-by-P array which represents P triangles on the head surface with each row representing one vertex of a triangle. The columns of this array are the X, Y and Z axis values respectively. This 3D array can be reshaped to a three-row matrix using matlab, each column of which represents one point on the head surface. The 2D cross-sections along sagittal plane are specified by the Z coordinate values. Each Z axis value represents one cross-section layer and can be separated using simple matlab routines. For each 3D data forty nine cross-sections were extracted in this manner.

Schwarz-Christoffel mapping for head cross-section data

The Schwarz-Christoffel (SC) mapping function [8] is a conformal mapping function of the complex canonical domain on to the physical domain. To adapt the SC function on the face plastic surgery design the complex canonical domain should be matched with the 2D head cross section data (s1, s2) obtained from the 3D head scanner and the physical domain should be matched with the unit disk as shown in the Figure 1. Lets consider the pre-surgery head cross sections of standard and target under the mapping and the SC mapping functions are $f_1(z)$ and $f_2(z)$ respectively where z is a point on the unit circle. Since the SC mapping is one to one, all the points on the head cross sections are mapped on to a unit disk. The points $w_1 = f_1(z)$, $w_2 = f_2(z)$,…, $w_n = f_n(z)$ represent the standard head cross section as well as the points $w_1 = f_1(z)$, $w_2 = f_2(z)$,…, $w_n = f_n(z)$, represent the target head cross section. The points $s_1 = g_1(z)$,…., $s_n = g_n(z)$ represent the standard head cross section after surgery. Next the conformal mapping for the target head data is obtained through $g_1(z)$. The relation between the polygon points z_1 ,…, z_n and the pre-vertices w_1 ,…, w_n can be described by the Schwarz-Christoffel formula as equation (1).

$$f(w) = f(w_0) + c \int_{w_0}^{w_n} \prod (\xi - w_j)^{\alpha_j - 1} \, d\xi \quad (1)$$

where, ξ is the dummy variable for integral.

After obtaining $g_1(z)$, $f1(z)$ and $f_2(z)$ from numerical analysis, we designed a method which can calculate the function $g_2(z)$ corresponding to subject's conformal mapping for that cross-section after surgery. It may be mentioned here that $g_1(z)$, $f_1(z)$ and $f_2(z)$ represent the post–surgery standard head conformal mapping for a desired cross-section, pre-surgery standard head cross section and subject's pre-surgery head cross section data respectively.

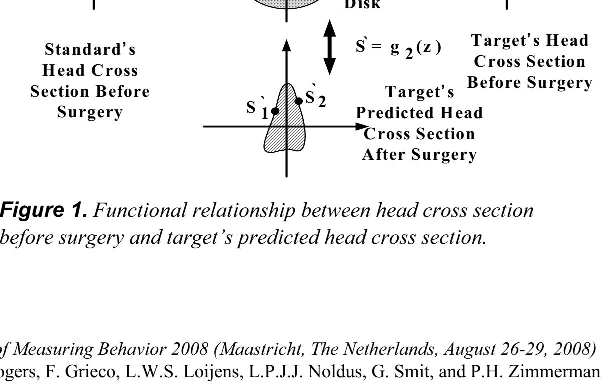


Figure 1. *Functional relationship between head cross section before surgery and target's predicted head cross section.*

By our proposed method we can design the face for plastic surgery for any kind of head using the following simple relation as shown in equation (2).

$$g_2(z) = \frac{g_1(z) \times f_2(z)}{f_1(z)} \quad (2)$$

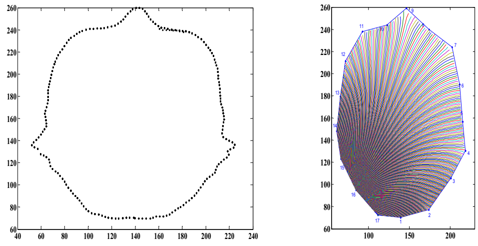
Results

To determine the conformal mapping function using Schwarz-Christoffel tool, the first step is to take the 2D head cross section data. Those data can be converted into complex data which is used to create a polygon and using those polygons the Schwarz-Christoffel map of that part can be obtained. This same procedure can be performed for a target head and the subject's head. Using the proposed method if Schwarz-Christoffel map of a cross section of a target head, its corresponding face plastic surgery design data and the subject's cross section data are given as input, the conformal mapping for subject's cross-section can be obtained.

Figure 2(a) shows the cross-section layer 18 of target's cross-sections. Figure 2(b) shows the polygon plot and conformal mapping for this layer. Figure 3 shows the pre and post surgery conformal maps for standard head data. Figure 4 shows calculated conformal map for target head after surgery using equation (2).

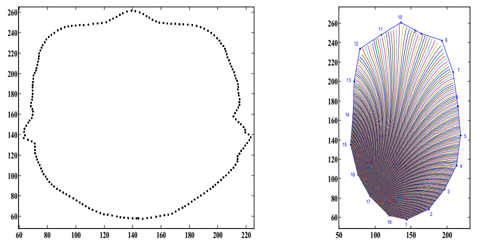
Conclusion

In this paper, a new approach for characterizing facial surgery is presented. A 3D head scanner is used to collect 3D head data and cross section data from distinct locations on the face for different individuals are taken. Thereby, a representation of the feature, more compact and effective for these applications is achieved. Further research is needed to assess the practicalities and its medical application in facial surgery.



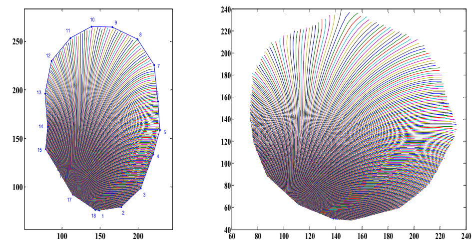
a *b*

Figure 2. *(a) Cross section for layer 18 of subject's head (b) Polygon Plot (blue) and Conformal Mapping of the cross-section.*



a *b*

Figure 3.*(a) Cross section for layer 18 of target head (b) Polygon Plot (blue) and Conformal Mapping of the cross-section.*



a *b*

Figure 4.*(a) Polygon Plot (blue) and Conformal Mapping of Subject's head cross section before surgery (b)Calculated Conformal Mapping of the cross-section for post-surgery.*

References

1. Waters, K., Terzopoulos D. (1991). Modeling and Animating Faces Using Scanned Data. *Journal of Vis. And Computer Animation* **2(4)**, 123-128.
2. Lee, Y., Terzopoulos D., Waters, K. (1995). Realistic Modeling for Facial Animations. *In Proc. ACM, SIGGRAPH*, 55-62. ACM Press.
3. Welch, W., Witkin, A. (1992). Variational surface modelling, *In Computer Graphics. Proc. ACM SIGGRAPH (1992)*, 157-166.
4. Altobelli, D.E., Kikinis, R., Mulliken J.B. (1993). Computer-assisted three-dimensional planning in craniofacial surgery. *Plast Reconstr Surg.* **92**, 576-585.
5. Berkowtz, S., Cuzz, J. (1977). Bio-stereometric analysis of surgically corrected abnormal faces. *Am. J. Ortho.* **72**, 526-538.
6. Yin, L., Hhs I.P. (1996).Constructing a 3D individualized head model from two orthogonal views. *Vis. Comput.* **12**, 254-267.
7. Kouchi, M., Mochimaru, M. (2005).Development of a Low Cost Head- Scanner for a Custom Face plastic surgery Making System. *Proc. of the 5th Symposium on Headwear Biomechanics*, 58-59.
8. Driscoll, T. A., Trefethen, L. N. (2002). Schwarz-Christoffel Mapping. Cambridge University Press.

Arthroscopic view: can we see more by quantifying the image quality?

G.J.M. Tuijthof, I.N. Sierevelt, and L. Blankevoort
Orthopaedic Research Center Amsterdam, Department of Orthopaedic Surgery, Academic Medical Centre, Amsterdam, The Netherlands, g.j.tuijthof@amc.uva.nl

Introduction

Minimally invasive surgery in orthopaedics (arthroscopy) has proven to be highly beneficial for patients as indicated by a decrease in morbidity and a rapid rehabilitation. Since the number of arthroscopic procedures continues to grow [3,7], reducing the operating time by optimizing the arthroscopic view would be desirable. The image quality during arthroscopic procedures is dependent on a number of factors: condition of the camera, the light source and the arthroscope, type of irrigation system, use of tourniquet, condition of the joint, portal placement, and skills of the surgeon. During arthroscopic surgery, the joint is continuously irrigated with saline fluid which reduces disturbances such as bleeding, air bubbles or debris. Even if the basic conditions are present, maintaining of a clear view is sometimes difficult [8]. To optimize the arthroscopic image quality, objective and quantitative measures are required, which at present are not available. Therefore, the goal is to develop a technique that enables objective and quantitative assessment of the arthroscopic image quality.

Approach, methods and results

The development of the technique was performed in three steps, two steps for defining arthroscopic image quality and one step for automatic detection.

Step 1

The arthroscopic image quality is primarily dependent on the presence of disturbances. A disturbance can be caused by different factors. But the only source of information is the movie of the arthroscopic view, which implies that only its effect on the view is determined. Seven types of disturbances were identified from prior observations of recorded arthroscopic knee procedures: Bleeding, Turbidity, Air Bubbles, Loose Fibrous Tissue, Attached Fibrous Tissue, Tissue too Close, and Instrument too Close (Figure 1). We propose to formulate descriptive definitions of disturbances in the arthroscopic view, which contain information on the size of the disturbed areas, and their duration. Initially, a percentage of 25% or more covering the image area was set as the threshold level for the presence of a disturbance. For the duration of a disturbance, the start time and end time were defined as the first and last image frame for which a disturbance conforms to the image area definition, respectively. With these definitions, the arthroscopic image quality was defined as good if no disturbances are present.

The unique interpretability of the definitions was evaluated with a time-action analysis, and tester agreement was assessed with the kappa statistic. A kappa value of 0.7 or more was considered to be a good agreement [1]. From ten routinely performed arthroscopic knee procedures, one minute was selected of each. The arthroscopic view was recorded in digital AVI-format (image resolution: 720 x 576, frame rate: 25 images per second). Four observers performed a time-action analysis with only the description of the disturbances at hand. For each separate disturbance, all arthroscopic images were marked as '1' if that particular disturbance was present, or marked '0' if that disturbance was absent. The kappa statistic was calculated (*Matlab*, version 7.0.4.365 (R14), The Mathworks). The main result is that the proposed descriptive definitions of disturbances show sufficient tester agreement (adjusted kappa >0.7) to qualify as uniquely interpretable. Only Turbidity showed a moderate agreement (kappa = 0.45).

Step 2

In the definition of disturbance, the required size of the disturbed area was defined arbitrarily.. A subsequent study was performed to determine quantitatively at which percentage the arthroscopic view becomes unacceptable based on expert opinions. The disturbances Tissue and Instrument too Close were discarded, as those were considered to be less relevant for clinical purposes, and the surgeon cannot always prevent them.

Thirty-two movie samples, 2 - 2.5 seconds in length, of the five disturbances were selected from videos of arthroscopic knee procedures. The movies showed disturbances covering different percentages of the image area, and were randomly presented. The average disturbance percentage of each movie was determined with *Matlab*, where in every fifth frame the disturbed area was indicated interactively. Twenty six orthopedic surgeons and thirteen residents were asked to indicate if the view was acceptable for each movie.

The most intolerable disturbance was bleeding. A clear transition from acceptable to unacceptable view was found for Bleeding (11% of covered area was acceptable; 25% not acceptable), and Air Bubbles (10% acceptable; 20% not acceptable) (Figure 2). Loose Fibrous Tissue showed a gradual transition where 25% was still accepted by a third of the surgeons. Turbidity and Attached Fibrous Tissue were tolerated up to 50% covered area by three quarter of the surgeons. Concluding, a safe value for an acceptable arthroscopic image can be set at 20% disturbed area.

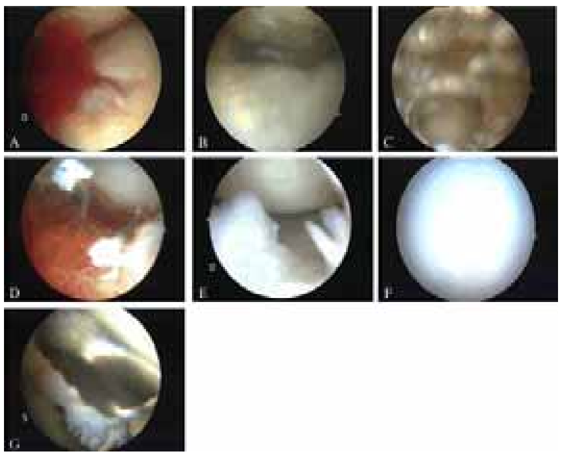


Figure 1. *Arthroscopic images represent the seven disturbances. A) Bleeding, B) Turbidity, C) Air Bubbles, D) Loose Fibrous Tissue, E) Attached Fibrous Tissue, F) Tissue too Close, and G) Instrument too Close.*

Step 3

The final step was to translate the results from step 1 and 2 into automated detection of the image quality. Initially, we focused on bleedings, as it was indicated as the most intolerable disturbance. As the background of the arthroscopic view is continuously changing, a generic feature representing blood had to be chosen for detection. A segmentation routine was designed (Matlab) that filtered a combination of Red, Green and Blue pixel levels representing the red tinctures of blood. Verification was performed by visual comparison of the segmented image with the original (Figure 3). The threshold level of 20% or more covering the image area was implemented.

Subsequently, the routine was used to analyze arthroscopic shoulder procedures, as these are known for the presence of bleedings. One surgeon performed ten shoulder arthroscopies. The preliminary results show that bleedings are less of a problem for arthroscopic rotator cuff repairs, 0% - 7% occurrence, but are frequently present in arthroscopic acromioplasty, 7% - 32% occurrence.

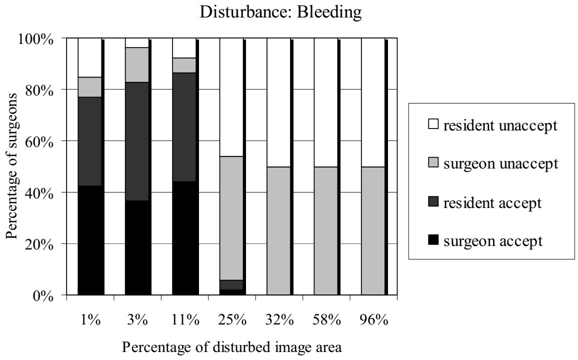


Figure 2. *Results of expert opinion on acceptability of bleeding covering a certain percentage of the image area.*

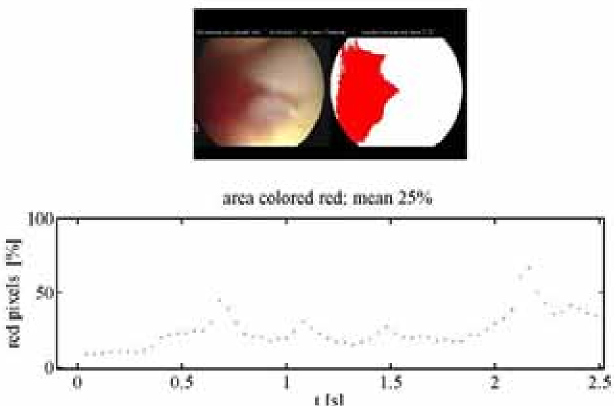


Figure 3. *Above: Automatically detected bleeding. Below: Percentage of red pixels in time for a short movie.*

Discussion

An objective and quantitative technique is developed for the analysis of arthroscopic procedures based on arthroscopic image quality. The approach can be used as a blue print for other endoscopic surgeries. For one disturbance automated detection was implemented. We aim at designing additional segmentation routines for Turbidity and Air Bubbles. The technique provides detailed information, which can be used to optimize surgical equipment such as the performance of irrigation pumps. A future direction is to develop this technique for performance monitoring of surgical skills.

References

1. Altman,D.G. (1991) Practical statistics for medical research, pp. 404. Chapman and Hall, London, England.
2. Ampat G, Bruguera J, and Copeland SA (1997) Aquaflo pump vs FMS 4 pump for shoulder arthroscopic surgery. *Annals of the Royal College of Surgeons of England* **79**: 341-344.
3. Baker, BS and Lubowitz, J. Meniscus injuries. eMedicine. 22-11-2006.
4. Dolk,T. and B.G.Augustini (1989) Three irrigation systems for motorized arthroscopic surgery: a comparative experimental and clinical study. *Arthroscopy* **5**: 307-314.
5. Muellner,T., W.A.Menth-Chiari, R.Reihsner, J.Eberhardsteiner, and L.Engebretsen (2001) Accuracy of pressure and flow capacities of four arthroscopic fluid management systems. *Arthroscopy* **17**: 760-764.
6. Ogilvie-Harris,D.J. and L.Weisleder (1995) Fluid pump systems for arthroscopy: a comparison of pressure control versus pressure and flow control. *Arthroscopy* **11**: 591-595.
7. Prismant. Ziekenhuisstatistiek - Verrichtingen. Internet: 2007. Research, Advise, Information and Education in Health Care, Utrecht. 15-8-2007.
8. Tuijthof,G.J., C.N.van Dijk, J.L.Herder, and P.V.Pistecky (2005) Clinically-driven approach to improve arthroscopic techniques. Knee.Surg.Sports Traumatol. *Arthroscopy* **13**: 48-54.

Using the Phantom Device for Rehabilition of the Arm in MS Patients: a Case Study

Karin Coninx[1], Chris Raymaekers[1], Joan De Boeck[1], Tom De Weyer[1], Geert Alders[2], Domien Gijbels[2], Bert Op 't Eijnde[2], and Peter Feys[2]

[1]Hasselt University, Expertise Centre for Digital Media and transnationale Universiteit Limburg, Wetenschapspark 2, 3590 Diepenbeek, Belgium, {karin.coninx, chris.raymaekers, joan.deboeck, tom.deweyer}@uhasselt.be

[2]REVAL Rehabilitation and Health Care Research Centre, Department of Health Care, University College of the Province of Limburg, Guffenslaan 39, 3500 Hasselt, Belgium, {galders, dgijbels, bopteijnde,pfeys}@mail.phl.be

Introduction

Multiple Sclerosis (MS) is a chronic progressive disease of the central nervous system. Depending on the distribution of lesions within the brain, MS may clinically present with impairments of strength, muscle tone, sensation, co-ordination, balance, bladder and bowel function, as well as visual and cognitive deficits, often leading to severe limitations of functioning in daily life. Studies of exercise therapy, focused on balance and walking outcome parameters, have shown a beneficial effect regarding muscle strength, exercise tolerance level, functional mobility and quality of life, while no important deleterious effects were reported [1]. Very few studies have properly investigated the therapeutic potential of arm training in persons with MS. Because training duration and training intensity are considered to be key factors for a successful neurological rehabilitation [2], we are investigating the value of robot-assisted rehabilitation of the upper extremity in persons with MS. More concretely, a virtual environment (VE) has been realized, which provides the patients with the training tasks to be carried out and monitors their progress and success rate. Not only visual feedback is presented in this VE; in order to train the patients with the execution of natural movements, a Phantom device is used, which generates force and thus proprioceptive feedback.

Below, we summarize the objectives of this research and we describe the system setup. Furthermore, we elaborate on the (multidisciplinary) research method being used to assess the potential of the Phantom as a training device in this context. Design and implementation of training tasks, and data logging, supporting measurement of the training effect as well as assessing usability, for the patients and the therapist are covered. We conclude with initial experiences of patients interacting with a proof of concept prototype environment.

Objectives and overall approach

The overall aim of our investigation is to assess the potential of the Phantom as a training device in the context of rehabilitation of MS patients with upper limb dysfunction. Besides measuring the effect of a Phantom-assisted training approach (Does the value of some quantifiable parameters such as muscle strength and arm function improve after repeated training?), we want to judge some usability issues such as acceptance of this kind of training program by MS patients.

In order to estimate the effect of the repeated Phantom-assisted tasks, it is important to design appropriate movement tasks (e.g. difficulty level allowing for effect measurements and allowing for as much carry over to actions in daily life as possible). Therefore, the team consists both of rehabilitation and computer scientists, working in a close cooperation with MS patients and clinical therapists of a rehabilitation centre. During the design and realization of the training tasks, it is important to pay attention to the possibility to measure the patient's behavior (e.g. by means of data logging with respect to movements with the device) and to "steer" the patient's behavior when performing the task by facilitating or obstructing the user's movements through the generation of appropriate forces by the Phantom device. This allows for personalization depending on the present capabilities of the patient and to change the training level.

A 3-week robotic training program is being set up, in which the patients perform the training tasks on a regular basis, i.e. approximately half an hour daily. At the beginning and at the end of this period, the user's arm motricity and functionality is tested using standardized neurological evaluation and arm function tests in order to judge if the patients' capabilities for executing manual tasks in daily life have been improved.

System setup

The training system consists of a standard PC and a Phantom device. In order to provide the patients with a large enough working field, a Phantom 1.5 is used in combination with a 19" monitor. This Phantom device is handled through a pen-like handle and allows tracking the handle's tip in 6 degrees-of-freedom and can generate forces in 3 degrees-of-freedom. The currently realized VE supports three training tasks (see figure 1).

In a first task, the patient sees a top-view of a road. A virtual car, steered by the Phantom device, has to be moved to from a start position via a various trajectory to an end position on the road. The patient is aided in this task by restricting the Phantom movement to the 2D plane in which the road is located. Furthermore a force is applied, which attracts the car to the centre of the road. The size of this force can be changed in order to change the difficulty of the task. Optionally, a viscosity force can be used to require a higher force from the patient, or to assist patients with tremor.

For the second task, users have to grab a book lying on a shelf

Figure 1. *Training tasks*

and place it in a bookcase. The forces used simulate the gravity and inertia of the book, which can be adjusted to the patient's capabilities. This task resembles natural object manipulation movements and requires a stable position to grasp the object as well as adaptive motor behavior towards the weight of the object.

The third task is a virtual implementation of the well-known plate tapping test, where patients need to tap two plates as often as possible within a given amount of time, or until 30 reciprocal tappings have been performed.

While performing these interactive tasks, the patients' movements are recorded at a rate of 200 Hz. Although, the analysis of the training data will be limited to the pre and post test, together with one intermediate test, the amount of data gathered is difficult to handle. Therefore, the movement data, the parameters of the tests (e.g. the forces used) and the patients' capabilities will be stored in a relational database. This allows us to filter the data according to specific research interests and to more easily compare different conditions. For instance, using simple queries on the database it is possible to only retrieve the data of patients with a tremor, horizontal trajectories of task 1 or to only compare successfully completed tasks.

User feedback

Up to now, a number of MS patients without or with only mild upper limb dysfunction have used the VE during a single session. The usability of handling the Phantom in the VE was evaluated with the 'System Usability Scale' and adapted questions scored using the Visual Analogue Scale. Subjects did not have difficulties with the visuospatial transformation neither with the force feedback, and commented positively on the VE. One MS patient, who has decreased motor coordination, had difficulties in stabilising her arm at the object in order to grasp the book, demonstrating that this task are likely challenging enough for patients with upper limb dysfunction.

Acknowledgements

Part of this research is financed by Interreg III (project 4-BMG-II=84) as well as the Province of Limburg (Belgium). The authors want to thank their consortium partners in this project: Department of Human Movement Sciences, Maastricht University, The Netherlands and Rehabilitation and MS Center, Overpelt, Belgium.

Part of the research at EDM is funded by EFRO (European Fund for Regional Development), the Flemish Government and the Flemish Interdisciplinary institute for Broadband technology (IBBT)

References

1. White LJ, Dressendorfer RH (2004) Exercise and multiple sclerosis. *Sports Med* **34**: 1077-1100.
2. Kwakkel G., van Peppen R., Wagenaar R.C., Wood D.S., Richards C., Ashburn A., Miller K., Lincoln N., Partridge C., Wellwood I., and Langhorne P. (2004) Effects of augmented exercise therapy time after stroke: a meta-analysis. *Stroke* **35**, 2529-2539.

Apolipoprotein E4 status is Associated with Agitated Behavior in Nursing Home Residents with Dementia

Diana Lynn Woods, Brittany Bushnell, and Haesook Kim
School of Nursing, University of California, Los Angeles, CA, United States, lwoods@sonnet.ucla.edu

Background

Multiple factors contribute to the development of behavioral symptoms in dementia (BSD). One such factor may be the apolipoprotein E (APOE) genotype. Although the role of APOE polymorphisms in the development of Alzheimer's disease (AD) has been studied extensively, the relationship of APOE and BSD has received only limited attention. The role of APOE genotype in moderating behavioral symptoms and psychopathology has been studied in community dwelling persons with AD, using caregiver proxy reporting, however, this relationship has not been studied in the nursing home population Information related to APOE genotype and the relationship to BSD would add needed data to assess risk factors for behavioral symptoms.

Behavioral Symptoms of Dementia are a significant clinical problem in nursing homes with a prevalence ranging between 30% and 84% [1, 2] While classifications and names may vary across studies, behaviors generally include four broad categories: physically aggressive, physically nonaggressive, verbally aggressive and verbally nonaggressive [3, 4]. The most frequent behaviors include problematic vocalization, restlessness, wandering, and pacing. In two prevalence studies of BSD, 38% of nursing home residents had repetitive purposeless activity, 29% had problematic vocalization, 18% acted inappropriately, and 9% wandered [5, 6]. Several studies have examined the association of resident characteristics such as mental status, dementia diagnosis, age and gender to behavioral symptoms of dementia [4, 7], with varying results. Given that the risk of AD is 47% in those persons exhibiting one ε4 allele rising to 91% in those exhibiting two ε4 alleles and BSD are highly prevalent, few studies have included genetic variability, as one of the resident characteristics that may be associated with the frequency and intensity of behavioral symptoms [8-10]

Several studies using proxy reporting have explored the relationship between APOE and BSD in community dwelling persons with dementia with mixed findings [9-12]. These recent studies indicate that the risk for BSD, such as restlessness, pacing and vocalization, may be increased relative to the APOE genotype. To date no studies have assessed this relationship using direct observation in nursing home residents with AD. The purpose of the study was to describe the association between BSD and ApoE genotype

Methods

This study employed a comparative non-randomized two-group design (BSD versus non-BSD group). The independent variable was the APOE genotype, obtained by buccal swabs. The dependent outcome variable was a cluster of six agitated behaviors (restlessness, escape restraint, tapping and banging, searching and wandering, pacing and walking, and vocalization) measured using the Modified Agitated Behavior Rating Scale (mABRS) [13]. Participants A comparative sample of residents, with and without behavioral symptoms, with moderate to severe dementia and who met inclusion criteria were selected from the 130 residents screened. Nursing staff screened residents for participation in the study using the Brief Agitated Rating Scale (BARS) screening tool [14]. Criteria for participation included: > 65 years; living on the unit for at least 2 months [15]; stabilization on anti-psychotic medication for at least 1 month; continued residency in the nursing home for the duration of the study; a diagnosis of dementia, according to the DSM-IV criteria [16], a Mini Mental State Exam (MMSE) [18] score < 21 (range 0-30), and for those residents who exhibited BSD, a score of >15 on the BARS [14]. Residents were excluded if they had acute psychiatric or physical illness; mental retardation or developmental delay, a stroke that caused gross motor, visual or sensory impairment, and end-stage or Parkinson's dementia (as reported in the participant's medical record).

Once proxy consent was obtained from families, 36 participants (21 with BSD and 15 with no BSD) were enrolled into the study, out of a total of 79 who met the criteria and were approached for consent. Since we were able to obtain genotyping on only 34 residents, only these residents are included in the analysis. Research Assistants observed and recorded resident behavior every 20 minutes between 8:00 am and 8:00 pm daily for 5 days. Buccal swabs were obtained for later APOE genotyping prior to the observation period.

Results

Of the 34 participants aged 71- 102 (M = 87.9, SD = 6.8), 86 % were female and 47% had at least one 4 allele. Participants had an average MMSE score of 10.43 of a total possible score of 30 (SD + 7.4, range 0 – 21), indicating moderate to severe dementia, and a mean of 3.92 medical co-morbidities (SD + 1.96, range 0 – 6). Twenty-three residents (63.9%) showed cardiovascular problems such as hypertension (36%), while 22% were diagnosed with depression. Sixty-nine percent of residents were prescribed analgesic, 36% an antipsychotic, 17% an anxiolytic, and 75% an antidepressant.

ANOVA for repeated measures was used to test the association between APOE4+/4- and behavioral symptoms. Mean behavior scores were grouped according to the presence or absence of APOE 4 (4+ = .45+.34), (4- = .24+.16). There was a significant difference in total behavioral symptoms, driven by restlessness and vocalization ($F_{1,30} = 4.40$, $p = .04$), between those who had an APOE ε4 allele (3/4, 4/4) present compared to those who did not. Odds ratio of 1.71 (95% CI, .42-6.96), supported these findings. Restlessness (manual manipulation) was significantly inversely correlated with MMSE score ($r = -.37$, $p = .03$) but not APOE genotype.

Conclusions: The 47% of participants with at least one allele is consistent with a diagnosis of AD. Findings indicate that the presence of the APOE4 allele increased the risk for behavioral symptoms in nursing home residents with dementia, and that those persons with a diagnosis of dementia and a lower mental status, reflected in the MMSE score, were more prone to restless behavior. Genotyping may prove useful to identify individuals at risk for behavioral symptoms.

References

1. Bedford S, Melzer D, Guralnik J.(2001). Problem behavior in the last year of life: prevalence, risks, and care receipt in older Americans. *J Am Geriatr Soc* **49**, 590-5.

2. Vance DE, Burgio LD, Roth DL, Stevens AB, Fairchild JK, Yurick A. (2003). Predictors of agitation in nursing home residents. *J Gerontol B Psychol Sci Soc Sci* **58**, 129-37.
3. Beck C, Frank L. Chumbler NR, O'Sullivan P, Vogelpohl TS, Rasin J. (1998). Correlates of disruptive behavior in severely cognitively impaired nursing home residents. *Gerontologist* **38**, 189-98.
4. Cohen-Mansfield J. Werner P. (1998). Predictors of aggressive behaviors: a longitudinal study in senior day care centers. *J Gerontol B Psychol Sci Soc Sci* **53**, 300-10.
5. Cariaga J, Burgio L, Flynn W, Martin D. (1991). A controlled study of disruptive vocalizations among geriatric residents in nursing homes. *J Am Geriatr Soc* **39**, 501-7.
6. Burgio LD, Park NS, Hardin JM, Sun F. (2007). A Longitudinal Examination of Agitation and Resident Characteristics in the Nursing Home. *Gerontologist* **47**, 642-649.
7. McCarty HJ, Roth DL, Goode KT, Owen JE, Harrell L, Donovan K. et al. (2000). Longitudinal Course of Behavioral Problems During Alzheimer's Disease: Linear Versus Curvilinear Patterns of Decline. *J Gerontol* **55**, M200-206.
8. Caselli RJ, Reiman EM, Hentz JG. Osborne D, Alexander GE. (2004). A distinctive interaction between chronic anxiety and problem solving in asymptomatic APOE e4 homozygotes. *J Neuropsychiatry Clin Neurosci* **16**, 320-9.
9. Craig, D., et al. (2004). Apolipoprotein E e4 allele influences aggressive behavior in Alzheimer's disease. *J Neurol Neurosurg Psychiatry* **75**, 1327-1330.
10. Scarmeas N, Brandt J, Albert M, Devanand DP, Marder K, Bell K. et al. (2002). Association between the APOE genotype and psychopathologic symptoms in Alzheimer's disease. *Neurology* **58**, 1182-8.
11. Gabryelewicz T, Religa D. Styczynska M. Peplonska B, Pfeffer A, Wasiak B. et al. (2002). Behavioural pathology in Alzheimer's disease with special reference to apolipoprotein E genotype. *Dement Geriatr Cogn Disord* **14**, 208-12.
12. Greenwood PM, Lambert C, Sunderland T, Parasuraman R. (2005). Effects of Apolipoprotein E Genotype on Spatial Attention, Working Memory, and Their Interaction in Healthy, Middle-Aged Adults: Results From the National Institute of Mental Health's BIOCARD Study. *Neuropsychology* **19**, 199-211.
13. Woods DL, Dimond M. (2002). The effect of therapeutic touch on agitated behavior and cortisol in persons with Alzheimer's disease. *Biol Res Nurs* **4**, 104-14.
14. Finkel SI, Lyons J, Anderson RL. (1993). A brief agitation rating scale (BARS) for nursing home elderly. *American Geriatrics Society* **41**, 50-52.
15. Brooke, V. (1989). Nursing home life: How elders adjust. *Geriatr Nurs* **10**, 66-8.
16. American, Psychiatric, and Association, Diagnostic and Statistical Manual of Mental Disorders (1994). *DSM-IV. 4th ed.* Washington, D.C: American Psychiatric Press.
17. Folstein MF, Folstein SE, & McHugh, P. R. (1975). Mini mental state: A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research* **12**, 189-198.

Eating rate in the treatment of eating disorders and obesity

P. Södersten, M. Zandian, I. Ioakimidis, and C. Bergh
Karolinska Institutet, Section of Applied Neuroendocrinology, Mandometer Clinic, Huddinge, Sweden, per.sodersten@neurotec.ki.se

The cause of eating disorders and obesity

Outcome of the treatment of anorexia and bulimia nervosa remains suboptimal [1], possibly because it is assumed that eating disorders are caused by anxiety disorders [2] although interventions targeting these disorders are not effective [3]. An alternative approach is based on the observation that eating disorder symptoms emerge in healthy people who volunteered [4] or were forced [5] to starve. Most patients have a history of reduced food intake; bulimic behavior, for example, is elicited by food deprivation [4,5]. Starvation changes eating behavior [4] and we hypothesize that the change mediates between physical state and cognitive state in the patients. Similarly, obesity may causally depend a change in eating behavior [6].

Experimental control of eating rate

Recording weight loss of a plate during meals yields a measure of cumulative food intake. Curves for eating rate displayed on a screen provide visual feedback during meals and subjects can follow these curves, because they see their eating rate appearing on the screen during meals. This method makes it possible to increase or decrease eating rate experimentally. Display of a scale, ranging from 0 - 100, on the screen allows simultaneous rating of satiety [7]. On the hypothesis that eating behavior is a cause of some aspects of eating disorders and obesity, this method is used to treat these conditions.

Linear eating: a risk for eating disorders and obesity

The cumulative food intake was fitted to a quadratic equation: $y = ax^2 + bx + c$, where a = change in the slope of the curve over time, b = initial rate of eating, and c = food intake at the start of the meal, i.e., 0. The cumulative satiety curve was fitted to a two-parameter sigmoid curve: $y = \alpha/(1+e^{-(x-x\ )/\beta})$, where $\alpha = 100$, i.e., the maximum of the satiety rating scale, β = steepness of the curve and x = time at which satiety has reached the half maximal value, i.e., the inflection point of the curve. These models fitted the experimental data satisfactorily; all square correlations were $r^2 \geq 0.99$. The averages of observed values are reported as medians; ranges and details of statistical analysis are omitted for clarity of presentation.

Subjects ate ordinary food (400kJ, 4.5 g protein, 18 g fat and 15 g carbohydrate/100 g) at lunchtime. Three weekly tests of 30 normal weight (body mass index, BMI = 22.2 kg/m^2), healthy men and women (aged 21.2) showed that intake and satiety is inter-individually stable. Forty women (similar BMI and age) were divided into linear; a = -0.1 (n=30) and decelerated; a = -2.4 (n=17) eaters in a control meal unassisted by visual feedback. Linear eaters ate at an initially lower rate (b = 29 g/min) than decelerated eaters (b = 44.2 g/min), but food intake (292 vs 288 g) and meal duration (10.9 vs 9.3 min) were similar. Satiety fitted the sigmoid curve in both groups; the inflexion point visible only among decelerated eaters, who reached a higher level of satiety (65.2) than linear eaters (51.4).

The women were challenged to eat 40% more food than they ate in the control test in the same period of time; this was achieved by asking them to follow the curve that they had generated in the control test and that was displayed on the screen. Decelerated eaters failed to adapt to the higher rate, ate less food (217 g) and reached a lower level of satiety (51.8). Linear eaters, however, ate more food (355 g) and reached about the same level of satiety (52.2). When challenged to eat 30% less food than they ate in the control test in the same period of time, the decelerated eaters ate about the same amount of food (274 g) and reached about the same level of satiety (64.2). Linear eaters, however, ate less food (249 g) and reached a higher level of satiety (60.2).

The results suggest that linear eaters, who eat at a constant rate, are unable to adjust their intake when challenged to eat at a rate that differs from their baseline rate. This is referred to as disinhibition, which is shown by those at risk for both eating disorders (too little food when eating slowly) and obesity (too much food when eating quickly). The default rate of eating may be decelerated, which may provide protection from the effect of disinhibition and therefore disordered eating.

Seventeen linear eaters (a = -0.2; b = 26.4), who overate by 16% in a test of disinhibition, practised eating at a decelerated rate three times/week during eight weeks by adapting to a decelerated curve; the women put food on their plate and an algorithm in the computer generated a decelerated curve with a = -1.7 and b = 53. After training, they ate about the same amount of food (261 vs 273 g) in about the same time (11.2 vs 12 min) as before training. However, they maintained a decelerated rate of eating (a = -3.9; b = 43) in the absence of visual feedback during the meal, and they did not overeat when challenged in a test of disinhibition (-6%).

These results suggest that linear eating, a possible risk for disordered eating, can be reduced by practicing eating at a decelerated rate [8], opening the possibility of using eating rate to treat eating disorders and obesity.

Treating eating disorders and obesity

A randomised controlled trial showed that the method described above had a significant effect on outcome, that an estimated 75% of 168 patients with anorexia, bulimia or unspecified eating disorder went into remission in on average 14 months and that 90% of 83 patents treated to remission remained free of symptoms in on average one year [9]. An interim analysis of an ongoing randomised controlled trial adapting the same principles to the treatment of childhood obesity showed that 31% of 36 patients lost > 0.5 BMI standard deviation scores (10-15 kg) compared to 11% among children treated as usual ($P < 0.001$).

References

1. Berkman, N.D., Lohr, K.N., Bulik, C.M. (2007). Outcomes of eating disorders: a systematic review of the literature. Int. *J. Eat. Disord.* **40**, 293-309.
2. Swinbourne, J.M., Touyz, S.W. (2007). The co-morbidity of eating disorders and anxiety disorders: a review. *Eur. Eat. Disord. Rev.* **15**, 253-274.
3. Holtkamp, K., Konrad , K., Kaiser, N., et al. (2005). A retrospective study of SSRI treatment in adolescent anorexia nervosa: insufficient evidence for efficacy. *J. Psychiatr. Res.* **39**, 303-310.

nervosa: insufficient evidence for efficacy. *J. Psychiatr. Res.* **39**, 303-310.
4. Keys, A., Brozek, J., Henschel, A., et al. (1950). The biology of human starvation. Minneapolis, MN: The University of Minnesota Press.
5. Burger GCE, Sandstead HR, Drummond JC. (1948). Starvation and malnutrition in Western Netherlands. Hague: The Hague General State Printing Office.
6. Bergh, C., Sabin, Shield J. et al. (2008). A framework for the treatment of obesity: early support. In Blass, E.M. (Ed.): Obesity: causes, mechanism and prevention. Sinauer: in press.
7. Zandian, M., Ioakimidis, I., Bergh, C., Brodin, U., Södersten, P. (2008). Effect of eating rate on food intake and satiety in women. *Physiol. Behav.* submitted for publication.
8. Zandian, M. , Ioakimidis, I., Bergh, C., Södersten, P. (2008). Reduction of dietary restraint, a risk for obesity. *Int. J. Obes*. submitted for publication.
9. Bergh, C., Brodin, U., Lindberg, G., Södersten, P. (2002). Randomized controlled trial of a treatment for anorexia and bulimi nevosa. *Proc Natl Acad Sci U.S.A*. **99**, 9486-9491.

Similarity maps of behavior sequences: Methods and software for pattern exploration and segmentation

Vicenç Quera
Department of Behavioral Science Methods, University of Barcelona, Spain, vquera@ub.edu

Researchers who study interaction are usually interested in detecting temporal patterns within observed sequences of behavioral codes. A pattern is "the repeated or regular way in which something happens or is done" (Collins COBUILD English dictionary); pattern is then synonymous with order and predictability, and the opposite of randomness and noise. Therefore, searching for patterns in sequences of codes that represent interaction amounts to detecting whether two or more codes often occur in succession, whether they tend to occur at the same time, or whether they tend to occur within a specific time intervals of each other. Different types of patterns can be potentially detected depending on how interaction codes are represented (as event sequences, timed event sequences, etc.; for a classification of types of sequential data, see [1]). A classical, quantitative approach to describing the sequence (e.g., [2]) consists in tallying transition frequencies among the codes and computing inferential statistics in order to know whether some transitions tend to occur more often than expected in case of randomness. However, besides that kind of molecular result, an exploration of the sequence as a whole can provide new insights as to whether certain patterns exist, where in the sequence they occur, and even whether they tend to repeat in different but comparable sequences. Such global exploration can be carried out by first computing indices of similarity among parts of the sequence (within specified time windows), then representing them as a two-dimensional map. Code repetitions, chains composed of certain codes occurring in succession, and possible patterns consisting of codes separated by a more or less constant interval may be detected by visual inspection of that map. Moreover, the map as a whole can indicate whether the sequence is probably random or whether it contains patterns, and, if so, where they tend to be located in the sequence.

Program RAP (for RAndom Projection, [3]) transforms event or timed event behavior sequences into such a map; both intra- and between sequences similarities can be computed. The program represents successive intervals or time windows in the sequence by quantitative vectors by means of an analytical technique called "random projection" [4], then computes the similarity between every time window and every other time window. Similarities, which are values ranging from 0 (no similarity between the two windows) to 1 (perfect similarity), are then displayed as grey pixels in an image, the greater the similarity the darker the pixel. Researchers can then visually inspect the image, and navigate it with the mouse cursor on a computer screen in order to explore its regions. These graphical representations are analogous to dot-plots (for exploring self-similarities in texts and in sequences of nucleic acids; e.g., [5], [6]), recurrence plots (for the analysis of continuous time series describing the behavior of dynamical systems; e.g., [7]), and waveform similarity plots (for the analysis of structure and classification of digital media streams; e.g., [8]).

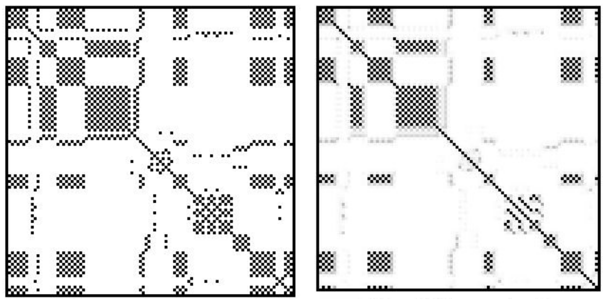


Figure 1. *Self-similarity maps created by program RAP for an event sequence of verbal interaction.*

Figure 1 shows two self-similarity maps created by RAP for an event sequence of verbal interaction in a couple. For the two maps, the sequence is represented top to bottom, and left to right; the map on the left was obtained by applying a moving time window containing one only code, and the one

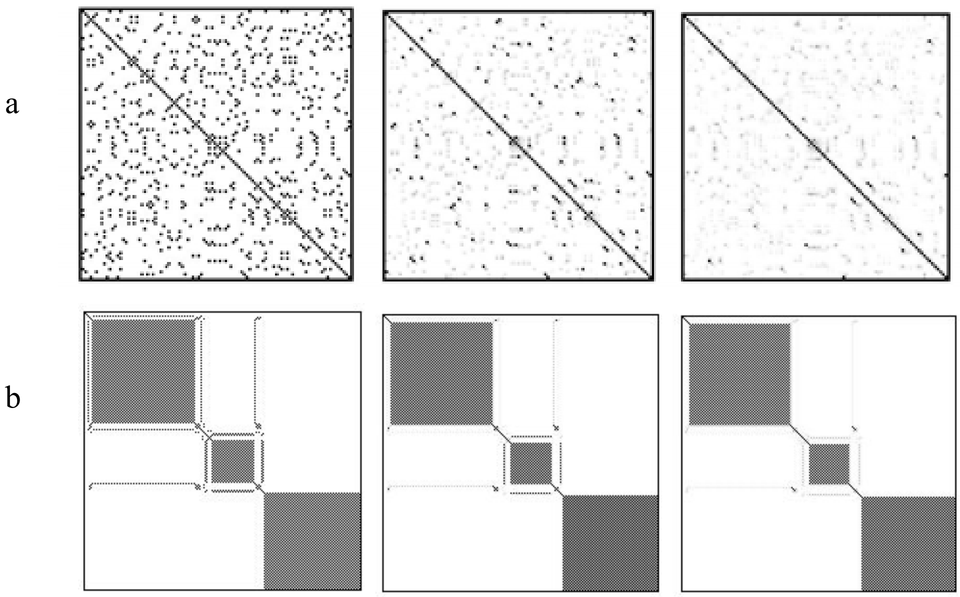


Figure 2. *Self-similarity maps for (a) a random event sequence, and (b) a highly patterned event sequence or verbal interaction.*

on the right was obtained by applying a window containing three consecutive codes (in the latter case successive windows overlapped). In both cases, main diagonals represent the similarity of every time window with itself, which is obviously maximum, and the images are symmetrical around their diagonals. Chequered regions indicate alternation of windows containing identical or very similar data; diagonal segments in the images, parallel to the main diagonal, indicate that certain chains of similar windows repeat in different parts of the sequence. Other hints about possible patterns are vertical and horizontal lines, either continuous or fragmented, which indicate that a certain code, or chain of codes, occurring in a position in the sequence repeats in several other positions. As a whole, a similarity map reveals general features of the sequences, and can be useful for classifying the interaction as patterned or random.

Figure 2a shows three self-similarity maps (for moving time windows containing 1, 2, and 3 codes, respectively) for an event sequence of verbal couple interaction in which husband responds to wife at random, and vice versa; Figure 2b shows three self-similarity maps for an event sequence of verbal interaction containing long runs of reciprocal interactions, which correspond to three big chequered squares along the main diagonal. For a random sequence, dark pixels in the similarity map tend to vanish rapidly as the width of the moving time window increases, while for a highly patterned sequence the proportion of dark pixels remains more stable. A self-similarity map can be further processed in order to reveal the temporal structure of the sequence; program RAP can detect segments in a sequence by correlating a Gaussian checkerboard filter along the map's main diagonal, and computing a novelty score (a technique proposed by [9]). Figure 3 shows a self-similarity map for a timed event sequence of mother-child interaction, obtained by applying a moving time window 20 s wide (successive windows overlapped by 10 s), and the resulting novelty score, whose peaks indicate segment boundaries, or temporal points at which significant changes are detected in the behavioral sequence.

Program RAP runs on Windows systems, reads sequence data files in *Sequential Data Interchange Standard* (SDIS) format (see [1]), and saves similarity maps and related graphical information as BMP images included in HTML documents. The program can be downloaded from www.ub.es/comporta/vquera

References

1. Bakeman, R., & Quera, V. (1995). *Analyzing interaction. Sequential analysis with SDIS and GSEQ*. New York: Cambridge University Press.
2. Bakeman, R., & Quera, V. (1995). Log-linear approaches to lag-sequential analysis when consecutive codes may and cannot repeat. *Psychological Bulletin* **118,** 272-284.
3. Quera, V. (2008). RAP: A computer program for exploring similarities in behavior sequences using random projections. *Behavior Research Methods* **40,** 21-32.
4. Mannila, H. & Seppänen, J. (2001). Recognizing similar situations from event sequences. *Proceedings of the First SIAM Conference on Data Mining*, Chicago. Available at: http://www.cs.helsinki.fi/~mannila/postscripts/mannilaseppanen siam.pdf
5. Church, K.W., & Helfman, J.I. (1993). Dotplot: A program for exploring self-similarity in millions of lines of text and code. *Journal of Computational and Graphical Statistics* **2**, 153-174.
6. Maizel, J.V., & Lenk, R.P. (1981). Enhanced graphic matrix analysis of nucleic acid and protein sequences. *Proceedings of the National Academy of Sciences* **78,** 7665-7669.
7. Eckmann, J.P, Kamphorst, S.O., & Ruelle, D. (1987). Recurrence plots of dynamical systems. *Europhysics Letters* **5,** 973-977.
8. Cooper, M., & Foote, J. (2002). Automatic music summarization via similarity analysis *Proceedings of the International Symposium on Music Information Retrieval*, 81-85.
9. Foote, J., & Cooper, M. (2003). Media segmentation using self-similarity decomposition. *Proceedings of SPIE,* **5021,** 167-175.

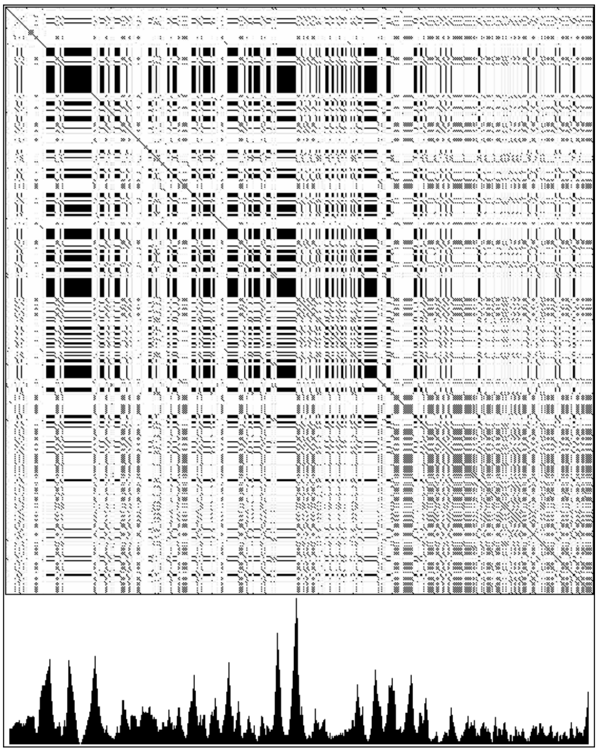


Figure 3. *Self-similarity map for a timed event sequence of mother-child interaction, and novelty scores indicating segment boundaries in the sequence.*

Analysis of the interaction contexts in soccer performed by the teams along the last 13 finals of the Eurocoup

Julen Castellano, Abigail Perea, and Lorea Alday
Department of Sports, University of the Basque Country, Vitoria, Spain, julen.castellano@ehu.es

A common characteristic in the Observational Methodology is the fact that behaviors can only be partly observed. Because of this fact, it becomes necessary to investigate if the observed variance is related with the individuals, the measuring tools, the place or other facets. The intention of this study is to find out which variables describe the game action in soccer and can be used to investigate its evolution.

Thirteen Championship final matches were observed in this investigation with a given taxonomic tool. Afterwards, a variance analysis of 4 facets (matches, area, result and interaction contexts) of the taxonomic tool was performed. Different techniques to estimate the variance components were used [1-4] to define which facets provide more information to the taxonomic tool and to determine which facets should be optimized. Furthermore, an investigation of the precision of the generalization has been performed with the objective of finding out in which way the facets have changed along the years.

The results confirm that the enchainment of the interaction contexts did not occur by chance, but according to an internal logic that could be described: certain offensive contexts are closer to scoring opportunities; meanwhile other contexts are more propitious for ball recovering or game construction. The estimated patterns did not change along the year, which gives us the possibility of obtaining very important information about the patterns in soccer.

References

1. Kane, M. (2002). Inferences about variance components and reliability-generalizability coefficients in the absence of random sampling. *Journal of Educational Measurement*, **39**, 165-181.
2. Marcoulides, G. A. (1996). Estimating variance components in generalizability theory: The covariance structure analysis approach. *Structural Equation Modeling*, **3**, 290-299.
3. Weiming, L., and Fang, Y. (2001). Model selections, variance component explanations and index comparisons in the application of generalizability theory: Comments on Ziu and Zhang (1998, 1999). *Acta Psychologica Sinica*, **33**, 467-470.
4. Wright, D., and Piersel, W. C. (1992). Components of variance in behavior ratings from parents and teachers. *Journal of Psychoeducational Assessment*, **10**, 310-318.

Polar Coordinate Analysis using Matlab, an application in soccer

Abigail Perea[1], Julen Castellano[1], M. Teresa Anguera[2], and Lorea Alday[1]
[1]*Department of Sports, University of the Basque Country, Vitoria, Spain, bigaper@hotmail.com*
[2]*Department of Behavioral Sciences Methodology, University of Barcelona, Spain*

The polar coordinate analysis [1] is a double data reduction strategy which provides a vector representation which determines the relation between the criterion behavior, this is, between the criterion which is our object of study and the other categories which conform the taxonomic system. With the intention of automatically perform the polar coordinates analysis a Matlab [2] script file was created.

Six matches played during the final phase of the Germany '06 World Championship were selected for this study, in which a taxonomic system (ad hoc) was defined with the objective of specifying the interaction contexts performed by the teams and the use of the space during the game. The observation and data registering was performed using SOCCAF v2.2video [3]. Afterwards, a sequential analysis using SDIS-GSEQ [4] and a polar coordinate analysis were performed.

The results obtained from the polar coordinates analysis allow describing the game action in soccer, adding the diachronic dimension of the events and combining the prospective and retrospective perspectives. By automatically performing the calculations using the Matlab approach, we have dramatically reduced the time for the calculations, as well as increased the quality of the results.

References

1. Sackett, G. P. (1980). Lag Sequential Analysis as a data Reduction Technique in Social Interaction Research. In D. B. Sawin, R. C. Hawkins, L. O. Walker y J. H. Penticuff (Eds.), *Exceptional infant. Psychosocial risks in infant-environment transactions*. New York: Brunner/Mazel, pp. 300-340.
2. Matlab. (1994-2005).The Math Works, Inc.
3. Perea Rodríguez, A. E., Alday Ezpeleta, L. and Castellano Paulis, J. (2006). Aplicación informática específica para la observación de la acción de juego en fútbol. In *Castellano Paulis, J.; Sautu Apellaniz, L. M.; Blanco Villaseñor, A.; Hernández Mendo, A.; Goñi Grandmontagne, A. y Martínez de Ilarduya, F. (coord.), Evaluación e intervención en el ámbito deportivo*. Diputación Foral de Álava, pp. 285-294.
4. Bakeman, R. and Quera, V. (1996). Análisis de la interacción. *Análisis secuencial con SDIS y GSEQ*. Madrid: RA-MA.

Identifying and analysing motor skills answers in the corporal expression and dance through OSMOS

Marta Castañer[1], Carlota Torrents[1], M. T. Anguera[2], and Maria Dinušová[1]
[1]*INEFC-Lleida, University of Lleida, Lleida, Spain, mcastaner@inefc.es*
[2]*Methodology of Behavioral Sciences Department, U B, Barcelona, Spain, tanguera@ub.es*

Abstract

The following paper outlines a new approach to the study of motor skills through the Observational System of Motor Skills –OSMOS- [1, 2]. The instrument is mixed and combines a field format with a category system for analysing natural contexts where body movement is relevant, as motor and sport practices, dance and psicomotricist situations. In the learning process of creative dance or similar disciplines, the application of the pedagogical model will be fundamental for the attainment of the proposed aims. During the practice, the instructions are usually open to generate a divergent production of motor skills answers. The paper also outlines a new approach to the analysis of time-based event records — in this case, for motor skills — known as T-pattern detection. Results suggest that it is possible to identify clearly all kind of motor skills based on the observational criteria and a further analysis of temporal behavioral patterns.

Key words

Observational System of Motor Skills, T- Patterns, Creativity, Corporal Expression, Dance, Pedagogical model

Objectives

To observe the production of motor skills answers in corporal expression and contemporary dance. To observe the capacity to generate motor skills answers in relation to the phases of all creative process: experimentation, fluidity, variability and elaboration.

To observe if "pedagogical models" based on exemplary instructions (the teacher gives a motor model as an example), metaphorical instructions (the teacher describes the tasks using a metaphor or guided vision) or descriptive instructions (the teacher explains the task using specific dance language) and "interindividual interaction" variables have any relevance to generate motor skills answers.

Methodology

120 sport university students without Dance experience participated in the study. A system of categories of observation settled down *ad hoc* to analyze the sessions using the Theme coder [3] to obtain analysis retardations. Later, we have used the Theme software for the detection of patterns of analyses was used that show to the relevance and the configuration to us of the registered events.

Instruments

The instrument (Table 1) is based on *changing criteria* (stability, locomotion, manipulation, *coreoespacial*[1] and temporal), each of which gives rise to a system of categories that are exhaustive and mutually exclusive. The data are then imported into the Theme software (www.patternvision.com) in order to detect hidden patterns.

Table 1. *OSMOS Observation Instrument. (Castañer, Torrents, Anguera, Dinušová, in press [2])*

CRITERIA	CATEGORIES
Stability	*Support stability* (Es): motor skills that enable body equilibrium to be maintained over one or several body support points, without producing locomotion (e.g. balancing actions)
	Stop stability (Ed): motor skills that enable the body to be projected by elevating it in space, without producing locomotion (e.g. jumps)
	Axial stability (Ea): motor skills that enable body axes and planes to be varied from a fixed point, without producing locomotion (e.g. turns)
Locomotion	*Propulsion-stop locomotion* (Lp): motor skills that occur at the start and finish of a body movement through space
	Sequential re-equilibrium locomotion (Ls): motor skills that enable a space to be moved through via the priority sequence of actions of the segments of the lower limbs (bipedestrian locomotion) or upper limbs (in inversion)
	Simultaneous coordinated locomotion (Lc): motor skills that enable a space to be moved through via the combined action of all body segments (e.g. quadrupedian locomotion)
Manipulation	*Impact manipulation* (Mi): motor skills in which certain body zones briefly come into contact with objects or other people
	Conduction manipulation (Mc): motor skills in which certain segments handle (for a given period of time) objects or other people
Coreoespacial	*Body changes* (C): evident variations in body posture and gestures
	Change in spatial direction (D): variations in the spatial direction of the movement
	Change of spatial level (N): change between the different spatial levels (low or floor work, middle or bipedestrian work, upper or aerial work).
	Combination of variations in body posture/gestures and spatial direction (CD)
	Combination of variations in body posture/gestures and spatial level (CN)
	Combination of variations in spatial level and direction (ND)
	Combination of variations in body posture/gestures, level and spatial direction (CND)
Temporal	*Time* (T): when there is a clearly observable change in the tempo of a motor action with respect to the previous one
Interaction	*Dyadic interaction* (Id): interaction with a partner
	Group interaction (Ig): interaction with more than one other group member

[1] *The body which generates these motor skills is located within a set of space-time coordinates which we have considered under a dimension termed coroespacial*

Results

Participants, when there was an exemplary model, generally produced significantly more different actions compared to the pedagogical model proposed by the teacher than exact reproductions, but they copied certain characteristics. Descriptive and metaphoric instructions seem to generate more varied answers than the exemplary. This is especially clear when using descriptive instructions, but on the other hand, metaphors help the students to understand the task and to be more motivated. Variations of the answers in all kind of tasks were significant in the category of time and body posture and gestures. Space changes or the variation of the type of interaction when there was a model were uncommon, and changes in the skill categories were specially evident in the tasks without model. It was also evident that the interaction with partners generates a different behavior compared to individual work, because participants try to agree with their partners. Moreover, the interaction seems to inspire the students and help them to feel themselves more creative.

References

1. Castañer, M., & Camerino, O. (2006). *Manifestaciones básicas de la motricidad*. Lleida: Publicacions de la Universitat de Lleida.
2. Castañer, M., Torrents, C, Anguera, M. T. & Dinûsová, M (in press): Instrumentos de observación ad hoc para el análisis de las acciones motrices en Danza Contemporánea, Expresión Corporal y Danza Contact-Improvisation. *Apunts*.
3. Magnusson, M.S. (2000). Discovering hidden time patterns in behavior: T-patterns and their detection. *Behavior Research Methods, Instruments, &Computers,* **32(1)**, 93-110.
4. Anguera, M.T. (2005). Microanalysis of T-patterns. Analysis of symmetry/asymmetry in social interaction. In L. Anolli, S. Duncan, M. Magnusson & G. Riva (Eds.) *The hidden structure of social interaction. From Genomics to Culture Patterns* (pp. 51-70). Amsterdam: IOS Press.
5. Jonsson, G.K., Anguera, M.T., Blanco-Villaseñor, A., Losada, J.L., Hernández-Mendo, A., Ardá, T., Camerino, O. & Castellano, J. (2006). Hidden patterns of play interaction in soccer using SOF-CODER. *Behavior Research Methods, Instruments & Computers,* **38(3)**, 372-381.
6. Pattern Vision, *Theme Coder* (software), 2001. Retrieved January 15, 2002, (from http://www.patternvision.com.)

Sequential pattern analysis software for educational event data

J. C. Nesbit, Y. Xu, P. H. Winne, and M. Zhou
Simon Fraser University, Burnaby, Canada, nesbit@sfu.ca

Many types of educational software are capable of writing event records to log files as learners interact with the software features. Such log files are potentially valuable to researchers investigating how learners access and manipulate information as they use the software. Educational researchers have explored several techniques for analyzing log files [1,4], especially in investigations of self-regulated learning [2,5,7].

gStudy is a software application for researching self-regulated learning [6] in solo and collaborative settings. Learners use gStudy to operate on multimedia documents packaged in learning kits. Because gStudy presents documents through a web browser, learning kits may contain media commonly found on the web such as text, video, diagrams, and animations. They can create notes, glossary entries, concept maps, HTML documents and other objects. For each session, gStudy writes a detailed log file in XML format. The gStudy log files contain a detailed list of time-stamped events tracing learners' behavior as they navigate through documents, tag information, and create information objects.

We have collected the log files of hundreds of university students who have used gStudy. In attempting to analyse these files we found that the behavioral unit or grain size that we wished to study, for example *making a note*, was at a higher level than the events actually recorded. An action that we wished to count might consist of multiple finer-grained events. Without a way to correct mismatching grain size, even counting the number of learner actions was difficult. Another challenge we encountered was characterizing commonalities in the sequential action patterns of a group of learners. Going beyond just counting the frequencies of actions, we sought to describe commonly occurring action sequences.

Data Mining Software for Educational Log Analysis

We developed analysis software to achieve two goals: (1) Identify coherent learner actions from a complex series of low-level events spaced along a time dimension, and (2) detect sequential patterns of those actions when focal actions are interspersed with unrelated actions.

To address the first goal, we developed a parser program called Log Validator. To prepare for parsing a log file with Log Validator, the researcher creates an *action library* that defines each multi-event action in terms of the fine-grained events in the log. For example, a make-a-note action might be defined to include these events: (a) select "make a new note" from the menu, (b) select the type of note, (c) input information into the fields of the note, and (d) close the note window. Log Validator matches events in a log file to the canonical action patterns defined by the researcher in the *action library* and builds an action file consisting of temporally ordered learner actions.

To address the second goal we developed a data mining program called Log Miner to detect which action subsequences are common and count all their instances. Figure 1 shows how a researcher would use the data analysis software to perform analyses consisting of the following five steps:

Customize an action library by defining actions of particular interest.

Use the action library and Log Validator to convert log files to more concise action files.

Use Log Miner to detect subsequences (patterns) that are common across the dataset or within groups.

Use Log Miner to identify and count the instances of common patterns detected in the previous step.

Use a 3rd-party statistical package (e.g., SPSS) to relate the occurrence of patterns to other variables.

The algorithm used by Log Validator to translate low-level events to higher-level actions is similar to that used in a computer language compiler. A common approach adopted in language compilation is the bottom-up parsing strategy, also known as shift-reduce parsing. It attempts to identify the most fundamental units first, and then to infer higher-order structures from lower ones. However, noise and extraneous events in logs of natural data would normally halt the shift-reduce parsing process because traditional compilers terminate with failure as soon as an error is detected in the input. We designed an adapted shift-reduce parser to allow parsing to continue in the presence of noise.

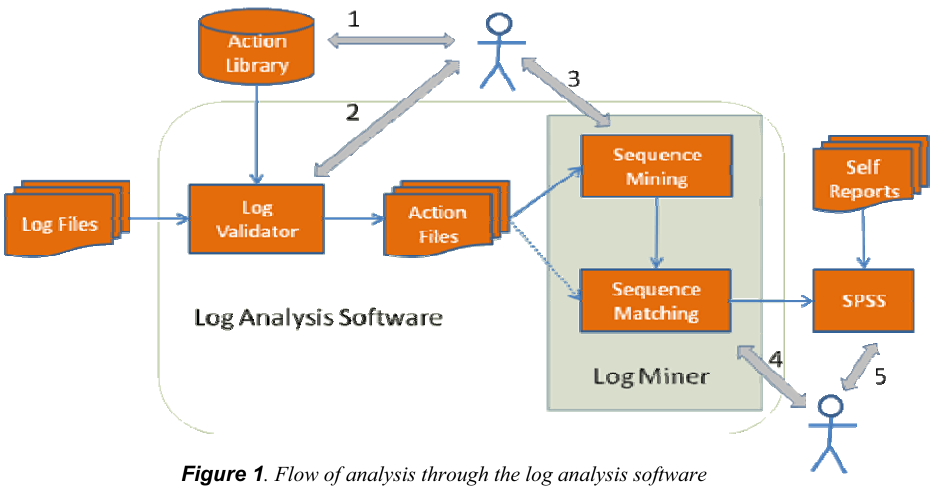


***Figure 1**. Flow of analysis through the log analysis software*

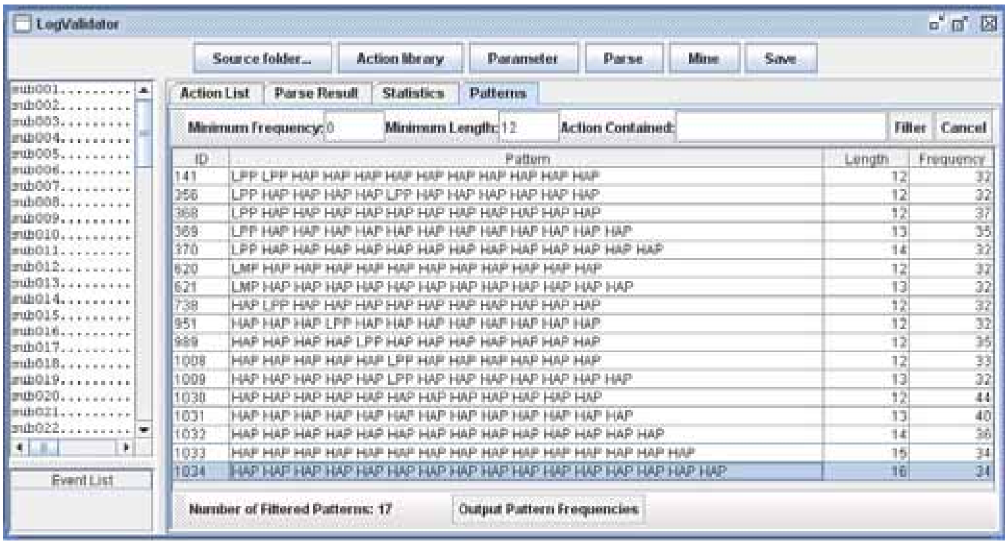


Figure 2*. The pattern mining algorithm returned hundreds of patterns occurring in at least 30% of 103 sessions. The longest common pattern (selected) is a series of 16 "HAP" actions.*

Sequential pattern mining [3] is a technique widely used in the data mining community for discovering common subsequences. For our purposes a sequence is defined as the ordered set of actions in an action files produced by Log Validator. A subsequence is an ordered set of actions within an action sequence. For example, *abcd* is a subsequence of the sequence *xaxxbxxcxxdx*. A subsequence is a new sequence which is formed from the original sequence by leaving out a number of the original actions without disturbing the relative positions of the remaining actions. Data mining algorithms find commonly occurring subsequences, called sequential patterns, which are subsequences that occur in more than a specified proportion of the input sequences.

In a large data set, it is usual for sequence mining to return several hundred common patterns, most of which have no relevance to the researchers' goals. Therefore it is necessary for the researcher to be able to search and filter common patterns so that relevant patterns can be identified. After discovering a large set of patterns, our analysis software can filter them according to length, frequency and content of the pattern. Finally, in a confirmatory pattern matching phase, the software finds and counts instances of the patterns in the current or subsequent data sets. The data from this phase can be imported to statistical software for further analysis. Figure 2 is a screen shot showing common patterns detected by the sequence mining component of our software and then filtered by pattern length.

Applications of the Log Analysis Software

We have used the log analysis software to analyse the studying behavior of 103 university students working with gStudy. Students had the opportunity to click on four different types of hyperlinks (e.g., HAP) and tag text with four different labels (e.g., LPP). The screenshot in Figure 2 shows that the longest common action subsequence consisted of repeatedly choosing to click on a particular type of hyperlink (HAP). We have also used the log analysis software to investigate the eye movements of 36 university students reading concept maps. At Measuring Behavior 2008, we will use these data sets to demonstrate applications of the log analysis software.

Conclusion

Pattern mining the data in log files is a promising approach to researching self-regulated learning. The method involves three stages. In the first stage the log data is organized into temporal sequences of theoretically meaningful actions. In the second stage, sequence modeling is applied to discover common patterns and theoretically relevant patterns are extracted. In the third stage, the occurrence of extracted patterns is counted to provide data suitable for statistical analysis.

References

1. Hadwin, A. F., Winne, P. H., & Nesbit, J. C. (2005). Annual review: Roles for software technologies in advancing research and theory in educational psychology. *British Journal of Educational Psychology* **75**, 1-24.
2. Hadwin, A.F., Nesbit, J. C., Code, J., Jamieson-Noel, D., & Winne, P.H. (2007). Examining trace data to explore self-regulated learning. *Metacognition and Learning* **2**, 107-124.
3. Agrawal, R. & Srikant, R. (1995). *Mining sequential patterns*, Proceedings of 11th International Conference on Data Engineering (ICDE).
4. Nesbit, J. C., & Hadwin, A. F. (2006). Methodological issues in educational psychology. In P. A. Alexander and P. H. Winne (Eds.). *Handbook of educational psychology* (2nd ed., pp. 825-847). Mahwah, NJ: Erlbaum.
5. Nesbit, J. C., Winne, P. H., Jamieson-Noel, D., Code, J., Zhou, M., MacAllister, K., Bratt, S., Wang, W., & Hadwin, A. F. (2006). Using cognitive tools in gStudy to investigate how study activities covary with achievement goals. *Journal of Educational Computing Research* **35**, 339-358.
6. Winne, P. H., Nesbit, J. C., Kumar, V., & Hadwin, A. F., Lajoie, S. P., Azevedo, R. A., & Perry, N. E. (2006). Supporting self-regulated learning with gStudy software: The Learning Kit Project. *Technology, Instruction, Cognition and Learning* **3**(1), 105-113.
7. Winne, P., Gupta, L., & Nesbit, J. C. (1994). Exploring individual differences in studying strategies using graph theoretic statistics. *Alberta Journal of Educational Research* **40**, 177-193.

Analysis of hidden patterns in team sports: Is there transversality in interaction contexts?

Javier Chaverri[1], Oleguer Camerino[1], and M. Teresa Anguera[2]
[1] *INEFC-Lleida, University of Lleida, Lleida, Spain, jchaverr@xtec.cat, ocamerino@inefc.es*
[2] *Department of Behavioural Sciences Methodology, University of Barcelona, Barcelona, Spain, mtanguera@gmail.com*

This study is part of a broader research project concerning various team sports and considers the interaction context as the core aspect that links team play.

In previous studies we have described the various instruments developed for soccer (SOF-5), handball (SOB-1), basketball (SOBL-2) and volleyball (SOV-1). All of these are combinations of field formats and category systems and they have been developed *ad hoc*, taking into account the characteristics of each one of them and their own rules; specifically, the use of field formats provides the instrument with dimensionality and flexibility, and makes it readily adaptable, whereas the category systems render it more consistent. In all these instruments a key role is played by the criterion corresponding to the *interaction context*, this being a core feature of the study of play due to its conceptual relevance, which is essential in terms of understanding better the mechanisms of interaction between opposing teams; here we analyse its potential transversality in the four sports studied, the obvious differences between them notwithstanding.

Objectives

To detect the patterns which enable us to understand the behavior of the interaction context in the four sports studied, and to conduct a microanalysis of this context.

Method

The observational design was nomothetic (because several teams are observed), point (because only one match is played between each pair of teams) and multidimensional (because there are various criteria in the respective observation instruments corresponding to football, handball, basketball and volleyball), and within-session monitoring was used. Ten matches for each of the studied sports were recorded. The recording instrument used was the *Match Vision Studio* software, which enabled us to code all the behaviors from the respective observation instruments indicated above. Using the *Theme 6* software (version beta) we conducted a microanalysis for the recordings of each of the four sports, studying patterns, sub-patterns and satellite behaviors, with special emphasis being placed on potential regularities between the various team sports studied and, above all, the behavior of the interaction context from a comparative perspective for these four team sports.

Results

The results obtained will be integrated into the broader study of which they form a part and will constitute the results of the analyses performed and those of the microanalysis regarding the respective interaction contexts.

Key words

T- Patterns, microanalysis, team sports, observation instrument, systematic observation.

References

1. Anguera, M.T. (2003). Observational Methods (General). In *Encyclopaedia of Psychological Assessment, Vol. 2* (edited by R. Fernández-Ballesteros), pp. 632-637. London: Sage.
2. Anguera, M.T. (2005). Microanalysis of T-patterns. Analysis of symmetry/asymmetry in social interaction. In *The hidden structure of social interaction. From Genomics to Culture Patterns* (edited by L. Anolli, S. Duncan, M. Magnusson & G. Riva), pp. 51-70. Amsterdam: IOS Press.
3. Anguera, M.T., Blanco-Villaseñor, A. & Losada, J.L. (2001). Diseños Observacionales, cuestión clave en el proceso de la metodología observacional. *Metodología de las Ciencias del Comportamiento*, **3(2)**, 135-161.
4. Anguera, M.T. & Jonsson, G.K. (2003). Detection of real-time patterns in sport: Interactions in football. *International Journal of Computer Science in Sport (e-Journal),* **2(2)**, 118-121.
5. Appleby, B. & Dawson, B. (2002). Video analysis of selected game activities in Australian rules football. *Journal of Science & Medicine in Sport,* **5(2)**, 129-142.
6. Assfalg, J., Bertini, M., Colombo, C., Del Bimbo, A. & Nunziati, W. (2003). Semantic annotation of soccer videos: automatic highlights identification. *Computer Vision and Image Understanding*, **92(2-3)**, 285-305.
7. Beauchamp, M.R., Bray, S.R., Fielding, A. & Eys, M.A. (2005). A multilevel investigation of the relationship between role ambiguity and role efficacy in sport. *Psychology of Sport and Exercise*, **6(3)**, 289-302.
8. Bloomfield, J., Jonsson, G.K., Polman, R., Houlahan, K., & O'Donoghue, P. (2005). Temporal patterns analysis and its applicability in soccer. In *The hidden structure of social interaction. From Genomics to Culture Patterns* (edited by L. Anolli, S. Duncan, M. Magnusson & G. Riva), pp. 237-251. Amsterdam: IOS Press.
9. Borrie, A., Jonsson, G.K., & Magnusson, M.S. (2001). Application of T-pattern detection and analysis in sports research. *Metodología de las Ciencias del Comportamiento*, **3(2)**, 215-226.
10. Borrie, A., Jonsson, G.K., & Magnusson, M.S. (2002). Temporal pattern analysis and its applicability in sport: An explanation and exemplar data. *Journal of Sports Sciences*, **20**, 845-852.
11. Castellano, J. (2000). *Observación y análisis de la acción de juego en el fútbol*. Unpublished Doctoral Thesis. Vitoria, Spain: Universidad del País Vasco.
12. Jonsson, G.K., Anguera, M.T., Blanco-Villaseñor, A., Losada, J.L., Hernández-Mendo, A., Ardá, T., Camerino, O. & Castellano, J. (2006). Hidden patterns of play interaction in soccer using SOF-CODER. *Behavior Research Methods, Instruments & Computers,* **38(3)**, 372-381.
13. Magnusson, M.S. (1996). Hidden real-time patterns in intra- and inter-individual behavior. *European Journal of Psychological Assessment,* **12(2)**, 112-123.
14. Magnusson, M.S. (2000). Discovering hidden time patterns in behavior: T-patterns and their detection. *Behavior Research Methods, Instruments, & Computers,* **32(1)**, 93-110.

Onboard video cameras and instruments to measure the flight behavior of birds

J.A. Gillies[1], M. Bacic[2], A.L.R. Thomas[1], and G.K. Taylor[1]
[1]*Department of Zoology, University of Oxford, Oxford, U.K., james.gillies@zoo.ox.ac.uk*
[2]*Department of Engineering Science, University of Oxford, Oxford, U.K.*

Summary

We have recently developed several novel techniques to measure flight kinematic parameters on free-flying birds of prey using onboard wireless video cameras and inertial measurement systems [1]. Work to date has involved captive trained raptors including a Steppe Eagle (*Aquila nipalensis*), Peregrine falcon (*Falco peregrinus*) and Gyrfalcon (*Falco rusticolus*). We aim to describe mathematically the dynamics of the relationship between the position, speed, acceleration and orientation of the bird's body and the position and orientation of its control surfaces. Such models allow us to answer specific questions about the operation of the flight control system, and to make quantitative comparisons between the control systems of different species. Although developed specifically for measuring the behavior of birds in flight, aspects of this method will be applicable to studies of any large free-ranging animal requiring onboard logging of serial data or wireless transmission of video data.

Inertial Measurement Unit, GPS, logger and harness

The birds carry an inertial measurement unit (IMU) and global positioning system (GPS) unit, together with up to 4 wireless video cameras. The IMU/GPS unit (MTi-G, Xsens B.V., Enschede, Netherlands) outputs 3-axis velocity, orientation, angular velocity and acceleration at 100Hz, together with latitude, longitude, altitude, and static pressure. The IMU/GPS unit and the GPS antenna are mounted dorsally on the birds using a removable harness with Velcro fittings, such that the unit is mounted directly above the bird's centre of mass. A serial datalogger (Antilog RS232, Anticyclone Systems, Morden, Surrey, UK) is mounted ventrally on the harness, and logs data from the IMU/GPS unit to a 1GB SD card (Kingston Digital, Fountain Valley, CA, USA). Power is provided by lightweight rechargeable Lithium Polymer cells. The wires connecting the IMU/GPS unit and logger are routed through one of the harness straps, while the controls for the IMU/GPS and logger are mounted ventrally on the harness to allow easy access whilst the bird is on its handler's fist. The harness itself is worn in a similar way to a backpack, with dorsal and ventral strips of webbing onto which the IMU/GPS unit and logger are attached. Three of the harness straps incorporate weak links, rated to cause the harness to detach in the event of snagging.

Cameras, transmitters, antennae, receivers and recorders

Miniature CCD 'snake' cameras (Misumi, Chung-Ho City, Taiwan) connected to miniature 2.4GHz transmitters (Misumi, Chung-Ho City, Taiwan) used to record and transmit PAL video footage of the wings, head and tail. Up to 4 cameras can be used simultaneously. Additionally a wireless video camera with mirrored image splitter (Synceros Inc., Ithaca, NY, USA) can be used to give a stereo pair of images in a single frame, allowing three-dimensional reconstruction of the surface of the wings or tail. Power is provided by lightweight rechargeable Lithium Polymer cells. Signals are received by 1m-diameter parabolic antennae, connected to customised receiver units (634-RX, Low Power Radio Solutions, Witney, Oxfordshire, UK). The analogue signal is recorded digitally to mini-DV tapes by a Sony Handycam DCR-HC40E camcorder (Sony Corporation, Minato, Japan). Sequences are downloaded to a PC using iMovie (version 3.0.3, Apple Incorporated, California, USA) and deinterlaced to provide video at 50fps using JES Deinterlacer (Jan E. Schotsman, 2004). To date, we have achieved a transmission range of >1km in line of sight.

Synchronisation of the IMU/GPS output with the video

In addition to the serial output to the data logger, the IMU/GPS unit also outputs a 1Hz square wave. This is broadcast with the video data over the otherwise-unused audio channel and recorded with the video stream to mini-DV tape. This enables the video sequences to be synchronised with output of the IMU/GPS unit.

Loading and aerodynamics

In order to observe normal flight behavior from the birds, the weight and aerodynamic drag of the instrumentation and cameras are kept as low as possible. The added load is always <5% of the body weight of the bird, which is significantly less than the weight of prey items that these species would routinely carry in the wild. In any case, the size of the species in question limits the number of equipment items that can be carried, and at present only the 2.5kg Steppe Eagle is large enough to carry all of the equipment items. The instrumentation and cameras mounted on the back of the bird are streamlined by covering them with a lightweight nacelle, while the logger is built in flush to the harness.

Video analysis

All video analysis is done using custom-written software in Matlab (The Mathworks Inc., Natick, MA, USA). Manual identification of natural features and automatic pattern or shape recognition software is used to follow the movements of the tail. This enables changes in the pitch, roll and spread angles of the tail to be monitored throughout a flight. Summary variables are also obtained for the motions of the wing tips.

System identification

System identification [2] gives us a route into understanding the flight control system of the bird species studied, by analysing the dynamical relationship between the control inputs it makes and the response of its body. System identification has been used successfully for over half a century to determine experimentally the dynamics of what in control engineering is termed the 'plant' of a control system—typically the physical system being controlled, and in our case the animal's flight dynamics. System identification is here used to fit a dynamical model describing the response of the bird's body to observed control inputs made by the tail. The gross model structure is specified using theoretical insights from flight mechanics and aerodynamics. Finer details of model structure and estimates of its parameters are identified using maximum likelihood methods, minimization of prediction error, or other similar optimization procedures. The end result is a predictive model explaining how a horizontal triangular tail is used to control and stabilize the flight of a bird, without the aid of a vertical tail fin as used by most conventional aircraft.

References

1. Taylor, G.K., Bacic, M., Bomphrey, R.J., Carruthers, A.C., Gillies, J.A., Walker, S.M., and Thomas, A.L.R. (2008). New experimental approaches to the biology of flight control systems. *Journal of Experimental Biology* **211**, 258-266.
2. Klein, V. and Morelli, E. A. (2006). *Aircraft system identification: theory and practice*. AIAA, Reston VA.

A Study for Identification and Behavioral Tracking of Honeybees in the Observation Hive Using Vector Quantization Method

Toshifumi Kimura[1] , Hidetoshi Ikeno[1], Ryuichi Okada[2], and Etsuro Ito[2]

[1]School of Human Science and Environment, University of Hyogo, Hyogo, Japan,
{kimura, ikeno}@shse.u-hyogo.ac.jp
[2]Kagawa School of Pharmaceutical Sciences, Tokushima Bunri University, Kagawa, Japan,
okd@ncp8.es.hokudai.ac.jp, eito@kph.bunri-u.ac.jp

Introduction

It has been taken an interest in colony and social behavior of animals in behavioral biology. Social insects, such as ant, bee and so on, have been studied as model systems for revealing functions and mechanisms of a crop of animals. It is widely known in the world that the honeybee has a kind of sociality since von Frisch et al. discovered the waggle dance [1].

In order to study the sociality of social insect, the researchers principally have to work hand-to-hand tracking individual behaviors in the experiments that Seeley et al. has done [2]. However, we can record honeybee behaviors in the observation hive by using COTS video as massive contents recently. As alternated, it is needed to take thousands of man-hours for analyzing vast amounts of records [3]. For an example, tracking of animal behavior is one of important but painful work for researcher. It must be very effective to provide a behavioral tracking support system not only individual but also social animals.

In this study, we proposed a new method for identification of individual honeybees and tracking their behaviors simultaneously for analyzing interactions among them. It has been developed a prototype system which is configured with four processes; 1) Separation of object and background by vector quantization, 2) Classification of candidate region of honeybee bodies, 3) Identification to honeybees and 4) Tracking honeybee behaviors. It was applied to the vide image data for a frame of observation hive. It was shown that 73% of individual (more than 500) in each video image are detected and tracked by our system.

Methods

There are several important and difficult points for automatically detection and tracking animal behavior from movie data. In our proposed method, process of behavioral tracking is consists of these four parts;

Separation of object and background by vector quantization

Classification to candidate regions of honeybee bodies

Identification to honeybees

Tracking of honeybee behaviors

Vector quantization method has been applied for various fields, image compression, signal processing, machine learning and so on [4]. It can be approximated each training vector by predetermined code vectors. We applied it for classification of objects, honeybee, nest and other things. In our preliminary analysis, it is shown that objects in the original image were classified into eight categories, and one of these is corresponded with the candidate of honeybee bodies. Therefore, it is possible to extract honeybee body regions by quantized the vector.

Figure 1a. Original Image

Figure 1b. Original Image

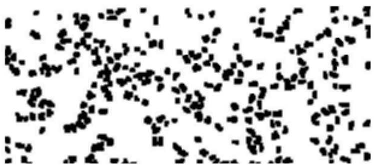


Figure 2a. Individual bee regions

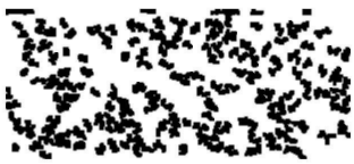


Figure 2b. Regions include plural honeybees

Figure 3. Identified honeybee

Behaviors of honeybees in the observation hive are recorded by handy camcorder as shown in Fig. 1 (a). As the result of first process, the candidates of honeybee bodies are detected as shown in Fig. 1(b). In the next step, each region is categorized by number of honeybee consist in it. After the twice application of image shrinking, object images are classifies into two kinds of region, based on their size. Fig. 2(a) and (b) show the results of classification in the single or plural honeybee body images.

Individual honeybee was identified and labeled based on decomposition of complex honeybee region images in the next step. In order to detect individual honeybee, an ID number is assigned on each honeybee extracted by the sequence of image processing. Video image is a time series data, so we indentified same honeybee in consideration of consistency of position. As the result of numbering, position of each honeybee on the frame of hive is calculated. In Fig.3, it is shown the result of numbered honeybee. In the last step, the frame by frame behavioral tracking is done by converting position of honeybee which has same ID number into time series data.

Results

We developed a prototype system using our proposed method (Fig. 4). As the result of application of our system to 387 images (30 frames per second), which are separated from the 10 minutes' video movie, it was extracted more than 73 % of individual honeybees (510 of 704, 522 of 718, 516 of 700). However, it was seen several failures in detection and identification.The major failures are mistakes of detection of honeybee images. The lost of object tend to lower the tracking precision. In our system, movie data separated into units consist of 50 images, then, trajectories of individual are tracked in each unit. By this treatment, behavioral traces of more than 400 honeybees obtained from movie.

Remarks

In this paper, we proposed a method to extract individual honeybees and track their movements. The system can extract more than 73% of worker honeybees in the hive. The prototype also show that it is possible to track the movement about a part of extracted bees, about 400 bees (about half of entire hive) in every 50 images. As a future work, we will develop a system that can track overlapped honeybees by analyzing behavioral causality. It could be are useful tools for behavioral analysis in order to revealing biological mechanisms of super-organized animals.

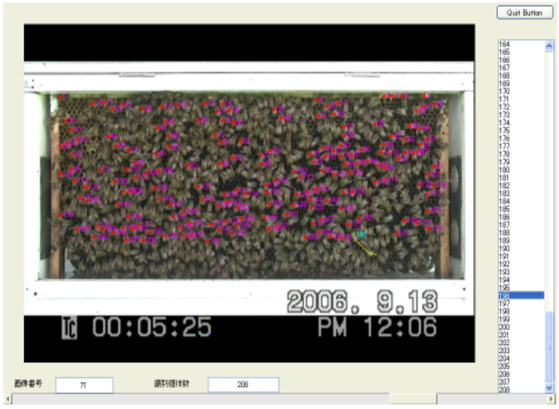


Figure 4. Prototype system

References

1. Frisch, K. v. (1967) The Dance Language and Orientation of Bees. Cambridge, Massachusetts : Harvard University Press
2. Seeley, T., Camazine, S., Sneyd, J. (1991) Collective decision-making in honey bees: how colonies choose among nectar sources, *Behavioral Ecology Sociobiology* **28**: 277-290
3. R.Okada, H.Ikeno, N.Sasayama, H.Aonuma, D.Kurabayashi, E.Ito (2008) The dance of the honeybee: how do honeybees dance to transfer food information effectively? Acta Biologica Hungarica, in press
4. Y.Linde, A.Buzo, R.M.Gray (1980) An Algorithm For Vector Quantizer Design, *IEEE Trans. Communications*, Vol. COM-**28:** 84-95

Performance characteristics of smaller-mammal GPS collars

Todd E. Dennis, Todd J. Landers, and Michael M. Walker
School of Biological Sciences, University of Auckland, Auckland, New Zealand, t.dennis@auckland.ac.nz

Tracking devices based on the world-wide, satellite-based NAVSTAR Global Positioning System (GPS) have been used extensively since the mid-1990's to acquire location and movement information from mammals for a wide variety of behavior-related research. Animal-borne GPS telemetry has been used to study aspects of movement [1] and activity patterns [2], migration [3], foraging behavior [4], and response to anthropogenic disturbance [5]. Compared with other means for collecting location data from animals roaming freely in their natural environments (such as direct observation, spoor and track surveys, bait and hair-trap stations, live-trapping grids, passage counters, and radio-telemetry) GPS telemetry provides several important advantages, including: 1.) high spatial and temporal accuracy; 2.) the capacity to collect large amounts of location information per individual over relatively long time periods; and importantly, 3.) the ability to determine location *in situ* without the need for human observers, thus avoiding collection [6] and observer-induced behavioral biases [7]. GPS telemetry permits cost-effective, continuous monitoring of individuals over a broad range of spatial and temporal scales even in challenging environmental conditions or in locations where human access is limited.

To date, tracking mammals with GPS-based devices has been limited almost exclusively to larger ground-dwelling or marine species, especially ungulates and carnivores. However, most mammals are much smaller than those previously studied using GPS telemetry, and thus are unable to bear the weight of such devices. As far as we are aware no studies have been published in the primary literature in which GPS telemetry has been employed on free-ranging mammals for any appreciable length of time (longer than one day) smaller than ~7 kg (a female Japanese macaque (*Macaca fuscata*) over a 9-day period [8]. Restriction of GPS telemetry to larger mammals has been due mainly to the difficulties and expense of engineering and constructing units light-enough (e.g., to meet the conventional < 5% of body-weight 'rule') and of appropriate physical dimensions to be carried by smaller species under field conditions for extended periods of time. Fortunately, many of the component technologies of GPS tracking devices have developed to the point that reductions of unit weight, size, and costs, as well as improvements in satellite acquisition rates, accuracy, and operational life now make it possible to deploy GPS tracking devices on much smaller species, greatly increasing the applicability of the method. The ability to track smaller mammals (and other animals) effectively offers solutions to many important biological questions and may prove to be one of the more powerful research techniques developed in the early 21st century (see [9]). Here, we report some performance characteristics of prototype GPS collars much smaller (≤ 105 g) than those used in previous studies. We tested the collars in natural conditions on a semi- arboreal marsupial, the brushtail possum (*Trichosaurus vulpecula*). Our study's general objective was to assess how well the collars could reconstruct the free-ranging movements of small-to-moderately sized mammals (> 2.5 kg) at fine spatial and temporal scales (see Figure 1). Specifically, we evaluate from field deployments the operational longevity of the collars at the scheduled sampling interval, the number of position fixes obtained, observation or 'fix-success' rates (ratio of the number of fixes acquired to the number attempted), and measures related to accuracy and precision of the estimated locations. In light of our findings we offer comments regarding the use of GPS tracking devices on smaller mammals in general.

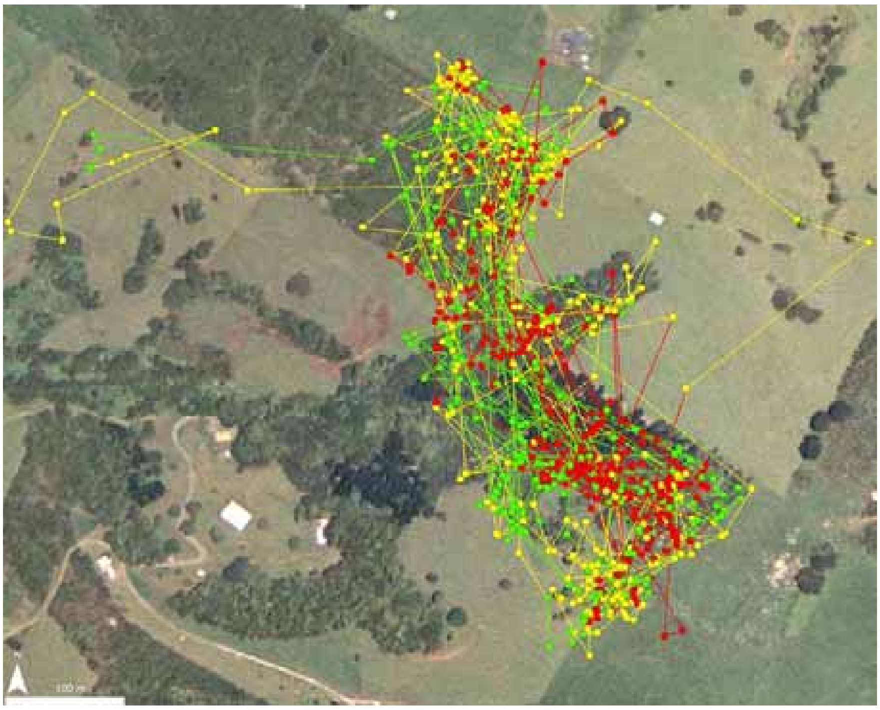


Figure 1. *Three weeks of GPS locations of a brushtail possum taken nightly at 15-min intervals. Colors refer to times of night green (early), yellow (middle), red (late). Scale bar in lower left corner is 100 m.*

References

1. Sheppard, J.K., Preen, A.R., Marsh, H., et al. (2006). Movement heterogeneity of dugongs, *Dugong dugon* (Muller), over large spatial scales. *Journal of Experimental Marine Biology and Ecology* **334 (1)**, 64-83.
2. Coulombe, M.L., Masse, A., Cote, S.D. (2006). Quantification and accuracy of activity data measured with VHF and GPS telemetry. *Wildlife Society Bulletin* **34 (1)**, 81-92.
3. Nelson, M.E., Mech, L.D., Frame, P.F. (2004). Tracking of white-tailed deer migration by global positioning system. *Journal of Mammalogy* **85 (3)**, 505-510.
4. Bishop C., Trathan, P., MacLean, G., et al. (2007). Novel GPS tags reveal fine-scale foraging movements of marine predators over extended time periods. *Comparative Biochemistry and Physiology A* **146 (4)**, S86-S86.
5. Rode, K.D., Farley, S.D., Robbins, C.T. (2006). Behavioral responses of brown bears mediate nutritional effects of experimentally introduced tourism. *Biological Conservation* **133 (1)**, 70-80.
6. Giuggioli, L., Abramson, G., Kenkre, V.M., et al. (2006). Theory of home range estimation from displacement measurements of animal populations. Journal of Theoretical Biology **240 (1)**, 126-135.
7. de Melo, L.F.B., Sábato, M.A.L., Magni, E.M.V., Young, R.J., Coelho, C.M. (2007). The secret lives of maned wolves *Chrysocyon brachyurus* (Illiger 1815) as revealed by GPS tracking collars. *Journal of Zoology* **271**, 27-36.
8. Sprague, D.S., Kabaya, M., Hagihara, K. (2004). Field testing a global positioning system (GPS) collar on a Japanese monkey: reliability of automatic GPS positioning in a Japanese forest. *Primates* **45(2)**, 151-154.
9. Wikelski, M., Kays, R.W., Kasdin, N.J., et al. (2007). Going wild: what a global small-animal tracking system could do for experimental biologists. *Journal of Experimental Biology* **210(2)**, 181-186.

Measuring aggression and threat-sensitive behavior in juvenile cod differing in size and nutritional state

L.A. Hansen[1,2], K. Skajaa[1,3], and B. Damsgård[1]

[1]*Norwegian Institute of Fisheries and Aquaculture Research, Tromsø, Norway, linda.hansen@fiskeriforskning.no*

[2]*Norwegian College of Fisheries Science, Tromsø, Norway*

[3]*University of Bergen, Bergen, Norway*

Absolute and relative differences in size and nutritional state may influence the social behavior and decision-making of individuals within a group. In the present study, individual feeding and swimming behaviors as well as aggressive interactions between two Atlantic cod, *Gadus morhua*, individuals of varying size and nutritional status were quantified. The aim of the study was to compare the relative influence of a change in nutritional status due to food deprivation (intrinsic factor), of absolute and relative size differences, and of nutritional state of the opponent (extrinsic factors) on aggression and threat-sensitive behavior in cod during the vulnerable juvenile stage. Aggressive behaviors, feeding incidence and swimming behaviors were scored and collected using the manual event recorder The Observer Video-Pro [1] and the automatic tracking system EthoVision [1].

The study was performed at the Tromsø Aquaculture Research Station (70° N), northern Norway. The behavioral effects of one to four days of food-deprivation and size differences (25 to 80 % relative weight difference) were investigated in pair-wise interactions with juvenile (0.7- 1.7 g) hatchery–reared Atlantic cod. The effects on behavior of food-deprivation in the smaller fish, food-deprivation of the larger opponent and the relative size differences were examined in thirteen different trial combinations where opponents varied in size and nutritional state. A total of 221 trials were conducted, testing 442 fish during the experimental period from 11 December to 23 December 2003.

Eight pair-wise trials were performed simultaneously in eight small aquaria (15.3 × 28.7 cm). Four cameras were mounted above the tanks, each recording fish in two aquaria. The walls and bottom of the aquaria were covered with non-reflective plastic film, reducing surface reflection. A white sheet surrounding the aquaria reduced the outside disturbance. The water inlet and outlet was on one short side of the aquaria, with a mesh preventing access to this area, and a plastic tube was positioned on the opposite side, extending down to 0.5 cm above the bottom from outside the white sheet. Feed could be delivered through the tube during the trials, keeping it in a specific area defined as the food zone of the trial arena in Ethovision. The food zone (15.3 x 10.0 cm) was one third of the trial arena. Prior to the trials the fish to be tested were carefully transferred to separate rearing tanks, and were not fed. Trials were run on four subsequent days allowing one or both of the contestants to be fed or food-deprived for either one, two, three or four days. Two fish, one large and one small, were introduced into each aquarium, and the fish were filmed continuously for 90 minutes. Food was supplied 60 minutes after trial initiation. At termination of the trial, weight and length was measured for each fish.

A fifteen-minute sub-sample of each trial recording was analysed, including the five minutes prior to supplying feed and the ten minutes after feeding. Individual aggressive and feeding behavior were manually scored using The Observer. Feeding and aggressive behaviors were recorded, pooling the aggressive behaviors charge, chase and bite [2], as were the receivers reactions to these behaviors (flee, hide, no reaction). The performer and receiver of aggressive behaviors were identified. Behavioral data on percentage of total available time spent in the food zone, frequency in the food zone, distance between the large and small fish, and swimming behavior were collected by use of EthoVision. Swimming behaviors include mean swimming speed, and time spent approaching and approach speed towards the small fish by the large fish.

The study suggests that large, dominant fish were consistently far more aggressive than smaller, subordinate fish, and foraged more frequently when food-deprived. Food deprivation did not increase the feeding frequency of small fish, but food-deprived small fish spent more time in the food zone than their larger, fed opponents. This suggests an increased interest in food and a change in threat-sensitive behavior. Food-deprivation of up to four days had no effect on the frequency of performed aggressive behavior and few effects on swimming behavior of either opponent. The results suggest that differences in size between opponents are a stronger influence on threat-sensitive behavior in juvenile cod than nutritional state. The benefit of risk-prone behavior in small, subordinate cod that has been food-deprived for up to four days is overruled by the cost of performing such behavior when the opponent is relatively larger.

References

1. Noldus Information Technology, Wageningen, The Netherlands
2. Petersson, E., Järvi, T., Olsén, H., Mayer, I. & Hedenskog, M. (1999). Male-male competition and female choice in brown trout. *Animal Behaviour* **57,** 777-783.

Measuring spontaneous vocal and facial emotion expressions in real world environments

Khiet P. Truong, Mark A. Neerincx, and David A. van Leeuwen
TNO Defence, Security and Safety, P.O. Box 23, 3769ZG, Soesterberg, The Netherlands
{khiet.truong, mark.neerincx, david.vanleeuwen}@tno.nl

Affective computing [1] has been introduced in the nineties as a research area that aims at designing systems that can recognize, interpret and synthesize emotions. For the development of these systems, databases containing recordings of natural emotions and descriptions (annotations) of these emotions are needed. In the past few years, there has been much discussion about the use of acted or real emotion data, and on how to annotate and measure emotion. In this paper, we elaborate on how to measure naturally occurring emotions in real world environments and how to establish a description (annotation) of these emotions with the aim to develop automatic emotion recognizers. We focus on vocal and facial expressions which can be measured in a relatively unobtrusive and simple manner. Learning from the findings and difficulties experienced during data collection in three different real world environments, we present a fourth experiment in which we measured and collected, natural vocal and facial emotion data in a multi-player gaming environment in a relatively short amount of time.

We had the opportunity to record vocal (and sometimes facial data) in three different real world situations where emotional expressions and behavior were expected:

1) Emergency situations on a naval ship which needed to be solved by operators.

In [2], an experiment is described in which the goal was to measure cognitive task load by processing several measurements, including vocal and facial expressions. In the ship control center, the operator had to deal correctly with the emergencies that occurred on the naval ship (e.g., fire or platform system failures). Several scenarios were designed to evoke low, medium or high task load with operators. High-quality webcams and head-mounted microphones were used to record video and audio. After each scenario, the participants had to rate task complexity and subjective effort on a five point scale. The idea was to find correlations between task load (or stress) and vocal and facial measurements. Unfortunately, the current vocal and facial analysis tools proved to be insufficiently robust against the realistic environmental conditions; e.g., background noise in audio and video, moving head postures etc., made it difficult to perform a reliable acoustic and facial analysis.

2) A flood (crisis) situation which needed to be dealt with by local authority members

This exercise was organized by the DECIS lab [3]. The 1-day scenario started with a simulated flood disaster in a small community in the Netherlands. During the crisis situation, a crisis policy team consisting of eight persons held five meetings in which the team members had to make time-pressured decisions. Vocal recordings were made using an 8-channel circular microphone array that was positioned at the table. We expected to find vocal expressions related to e.g., stress or frustration. However, this seemed less apparent than expected: the number of emotional expressions found in the meetings was too low to perform vocal and statistical analysis. An alternative way to describe this type of data could be to apply discourse analysis which can give more insight in the interaction between team members.

3) Players immersed in a virtual reality game

Exercise in Immersion (developed by Marnix de Nijs and V2_lab [4]) is an art-game played in an existing physical space. The player(s) wore a head-mounted display and a crash suit and immersed themselves in a virtual world. We expected a lot of vocal expressions, due to the player's immersion in the game and the "fun" factor. High quality voice recordings were made through close-talk microphones. Players were asked to "think aloud". Afterwards, the players filled in two questionnaires that were related to the emotions felt and the amount of presence experienced. For measuring presence, we used the Igroup Presence Questionnaire [5]. We are currently analyzing the data, and we can already observe that the "think aloud" procedure might not be very suitable for voice analysis since the players reported that it felt awkward to "think aloud" during the game which might have affected the way they express their emotions vocally and the naturalness of the speech.

Taking into account the "lessons learned" from the three previous recording sessions (environmental "noise" should be reduced as much as possible, the sparseness and naturalness of emotion expressions should be dealt with), we designed a fourth experiment to acquire annotated, multimodal, natural emotion data in a relatively short amount of time [6]. Participants played a first person shooter video game (Unreal Tournament) against each other in teams of 2x2. In total, we recorded 1120 minutes audiovisual data (28 participants x 20 minutes x 2 sessions). Several steps were taken to evoke as much vocal (and facial) interaction and expressions as possible by 1) asking participants to bring a friend as a team member, 2) manually generating "surprising events" in the game (e.g., the sudden appearance of monsters), and 3) granting bonuses for the winning team and good team collaboration. High quality audio recordings were made through close-talk microphones and recordings of the face were made with high quality webcams (placed at eye level). To obtain emotion annotations in a relatively short amount of time, we decided to use self-reported emotion values as subjective emotion measurements. After playing the video game, the participants had to watch their own video twice, and had to annotate their own emotions in two different ways: 1) choose an appropriate emotion label from a given set of labels whenever applicable (event-based, selection of multiple emotions at the same time was possible), and 2) give an arousal and valence estimation on a scale of 0-100 (continuously, each 10 seconds).

Frustration seems to be the most frequent emotion that was reported by the participants themselves (see Figure 1). Figure 2 presents the self-reported arousal and valence values: we can observe that high arousal is often accompanied with high valence. These self-reported emotion measurements need further investigation: we have to deal with both the personality of the participant and the subjectivity of the annotator (which in this case is equivalent to the participant).

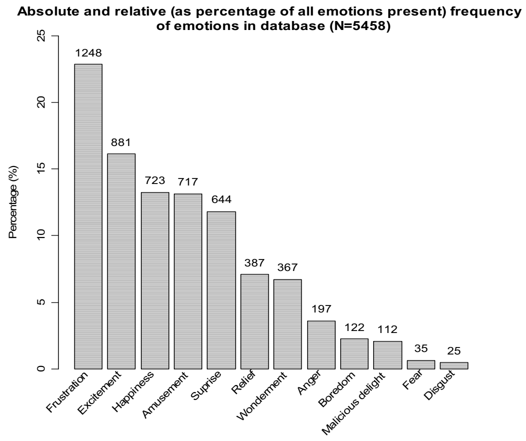


Figure 6. *Categorical emotion labeling*

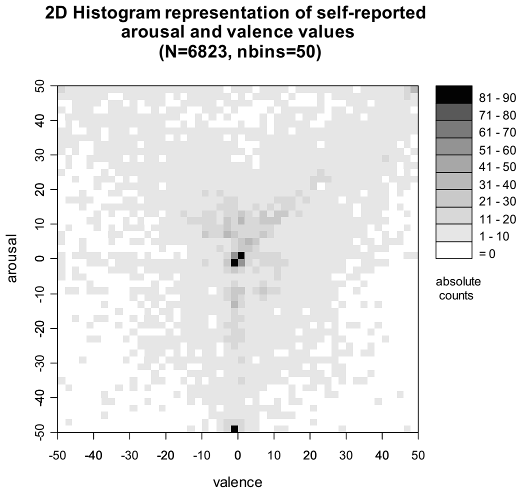


Figure 5. *Continuous arousal and valence labeling (darker areas indicate higher absolute frequency)*

To summarize, we can identify two main challenges for building automatic emotion recognizers: 1) the development of databases containing natural emotions that are annotated and measured in a reliable way, and 2) the development of robust automatic emotion recognizers for real-time emotion sensing in the real world. Our gaming experiment seems to provide a solution for the first problem (see Figure 3 for some example screenshots). However, we still need to validate our subjective emotion annotations which can be done by applying the "theory on context effects" (introducing a stressor in the game evokes a stress reaction) or by using multiple annotators (high agreement between the annotators can indicate high reliability). The second challenge also implies that emotion recognizers should be able to deal with "gradations" of emotions (which are very common in realistic emotions) instead of only the extremes. In future research, we will take up these challenges and measure emotions in complex, hectic environments (e.g., process control) with the aim to develop real-world emotion sensing systems.

Acknowledgements

This study is supported by MultimediaN, a Dutch BSIK-project (http://www.multimedian.nl).

References

1. Picard, R.W. (1997). *Affective computing*. MIT Press, Cambridge, MA.
2. Grootjen, M., Neerincx, M.A., Weert, J.C.M. van, and Truong, K.P. (2007). Measuring cognitive task load on a naval ship: Implications of a real world environment. *Proceedings of 12th HCI International (HCII) 2007, Beijing, China, July 22-27 (LNAI 4565),* 147-156.
3. http://www.decis.nl *[website]*
4. http://www.v2.nl *[website]*
5. http://www.igroup.org/projects/ipq/ *[website]*
6. Merkx, P.A.B., Truong, K.P. and Neerincx, M.A. (2007). Inducing and measuring emotion through a multiplayer first-person shooter computer game. *Proceedings of the Computer Games Workshop 2007, Amsterdam, The Netherlands*, 231-242.
7. http://www.noldus.com/site/doc200705001 *[website]*
8. Uyl, M. den; Kuilenberg, H. van (2005). The FaceReader: Online Facial Expression Recognition. *Proceedings of Measuring Behavior 2005,* 598-590.

Figure 7. *Screenshots of facial expressions recorded in gaming experiment, emotion classification provided by FaceReader [7,8]*

The pattern of Facial Expression among Iranian Children

Asghar Dadkhah, Moloud pourmohammadi, and Peymaneh Shirinbayan
Research Institute, University of social welfare and rehabilitation, Tehran, Iran, asgaredu@uswr.ac.ir

It is largely acknowledged that the ability to experience and express emotions is a basic human capacity, and studies concerning emotion lexicon and recognition of facial expression have strengthened theories of universality of human emotional potential, especially concerning the experience of so-called 'basic emotions' such as happiness, anger, sadness, and fear [1]. Longstanding theoretical perspectives on cultural differences present Western cultures as more individualistic than Middle Eastern and Asian cultures [2,3,4]. Values, norms, and beliefs of Western countries are often characterised as reflecting traditions that stress the importance of independence and autonomy of the self [5]. In Middle Eastern countries such as Iran, however, the self is viewed as more committed to a group, and values like group harmony and cohesiveness are highly emphasised [6], giving rise to what Gregg (2005) calls 'social selves'. Social values such as loyalty, respect, and hierarchy in a group are critical in the more collectivisitic-orientated cultures [7]. Markus and Kitayama [8] have conceptualised this focus on connectedness among individuals in terms of an 'interdependent' construal of the self.

The purpose of the present study was to understand and compare whether Iranian children enable to identify common emotions? And how do they express their emotions Specifically on Facial Expression? In the study reported below, we use a structured interview to identify 10- to 11-year children's self-reported use of emotional display rules in their *own* experience. This emic approach to the topic is likely to help us draw more culturally sensitive conclusions about children's emotional development, in comparison with an etic approach based on using hypothetical vignettes that are themselves potentially culturally biased. We have chosen to focus on children aged 10-11 years because developmental research suggests that children's capacity to reflect on their emotions and behaviors, in relation to various internal states, improves significantly during the primary school years.

Participants included 56 Iranian children, with 28 boys and 28 girls. The children were aged 10-11 years (mean age in months= 125.54, *SD* = 5.68) and were recruited through primary schools. Children's self-reported use of emotional display rules was assessed through a structured interview [9], which was administrated in Persian. The interview was translated into Persian and backward translated. Three key questions were asked:

1. Did you ever feel angry, but you did not want other people to know?
2. When was it that you did not want other people to know you were angry?
3. Why did you not want other people to know how you felt?
4. The same questions were repeated for sadness, joy and fear.

Children's responses were tape-recorded, transcribed, and all transcripts were double-coded by two people. The inter-rater reliability was assessed by means of Cohen's kappa for each question, separately for each emotion. Kappa varied from .79 (question 3; sadness) to .98 (question 1; anger). Differences between the coders were resolved by discussion. The coding procedures reported below were undertaken separately for each emotion.

The analysis of the data indicated that:1) all children were enabled to identify and differentiate all four emotions from each other,2)most of students Hide their happiness, anger, fear and sadness,3)they hide their emotions in Specific situation such as school and home,4) they hide happiness and anger against peers and hide fear and sadness in front of adults,5) children hide all four emotions through Neutralization,6) they appeared neutralization through Facial Expression in happiness , anger and fear, and Behavioral Expression in sadness,7)Pro-socially was the most important reason to hide happiness and anger and, Self-protection was the essential cause to hide fear and sadness,8) they were enabled to hide all four emotions Successfully,9)there was not any significant differences between girls and boys students on the basis of all eight criteria: Identification of Emotions Hide, Situation, Public, How, Expression, Why and Success. The present study provides the basic knowledge about Iranian children emotional expression which can be used in cultural, educational and therapeutic contexts.

Iranian children reported having experienced situations where they concealed each of the four emotions. Furthermore, there were systematic variations in the references for emotional display rules (see Table 1). The children referred more often to family audiences and less often to peer audiences, and more often to pro-social motives and less often to self-protective motives. Further analysis revealed that children were able to identify pro-social and self-protective motives for concealing emotions from a family audience (especially pro-social motives for concealing happiness and sadness, and self-protective motives for concealing fear and sadness), but were significantly less likely to identify self-protective motives for concealing all three negative emotions from a peer audience. The identification of display rule examples involving family members may relate to the heightened salience of status, hierarchy, and authority within Iranian family situations (Beeman, 2001); maintaining social order in these kinds of settings, through regulation of emotional expressions, may be a critical focus of socialization in this culture.

References

1. Ekman, P., Friesen, W. V., & Ellsworth P. (1972). Conceptual Ambiguities. In P. Ekman (Ed.). *Emotion in the human face: Studies in emotion and social Interaction* (p. 7-21). Cambridge: Cambridge University Press.
2. Gregg, G. S. (2005). *The Middle East: A cultural psychology.* Oxford: Oxford University Press.
3. Markus, H. R., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review* **98**, 224-253.
4. Triandis. H. C. (1995). *Individualism and collectivism.* Boulder, CO: Westview Press.
5. Zahn-Waxler, C., Friedman, R. J., Cole, P. M., Mizuta, I., & Hiruma, N. (1996). Japanese and United States preschool children's responses to conflict and distress. *Child Development* **67**, 2462-2477.
6. Diener, M. L., & Lucas, R. E. (2004). Adults' desires for children's emotions across 48 countries: Associations with individual and national characteristics. *Journal of Cross-Cultural Psychology* **35**, 525-547.

7. Cole, P. M., Bruschi, C. J., & Tamang, B. L. (2002). Cultural differences in children's emotional reactions to difficult situations. *Child Development* **73**, 983-996.
8. Markus, H. R., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review* **98**, 224-253.
9. Rieffe, C, Potharst, E, Meerum Terwogt, M, Begeer, S, Stockmann, L. & Cowan, R. (2005). Expressie van emoties bij hoogfunctionerende kinderen met autisme [Emotion expression of high-functioning children with autism]. *Kind en Adolescent [Child and Adolescent]* **14**, 1-10.

Table 1. *Means (and standard deviations) of Iranian children reporting display rule use, subdivided by Emotion, Motive, and Audience*

Emotion	Audience	Motive	Iranian
Happiness	Family	Pro-social	.24 (.43)
		Self-protective	.02 (.16)
	Peer	Pro-social	.22 (.42)
		Self-protective	.15 (.36)
Anger	Family	Pro-social	.08 (.27)
		Self-protective	.08 (.27)
	Peer	Pro-social	.33 (.48)
		Self-protective	.12 (.33)
Fear	Family	Pro-social	.08 (.28)
		Self-protective	.42 (.33)
	Peer	Pro-social	.00 (.00)
		Self-protective	.27 (.45)
Sadness	Family	Pro-social	.22 (.42)
		Self-protective	.20 (.40)
	Peer	Pro-social	.09 (.29)
		Self-protective	.09 (.29)

Using Noldus' The Observer XT to analyze videotaped footage of Deaf signers learning to read with a tutor

D.A. Ducharme and I. Arcand
School of Rehabilitation Sciences, Faculty of Health Sciences, University of Ottawa, Ottawa, Canada, dducharm@uottawa.ca, iarca001@uottawa.ca

Despite years of research on reading problems in deaf students, we still do not know how deaf signers who read well actually crack the code of print [1]. Three factors have recently been discussed in the literature as having an impact on the reading outcomes of deaf readers. First, the use of specific signing skills, namely fingerspelling, initialized signs and chaining has been suggested as a way to create associations between sign and print. Second, it has been suggested that deaf teachers and hearing teachers make different use of these skills. Third, students who are taught by deaf and hearing teachers seem to also use different reading strategies that reflect the mode of teaching to which they are exposed. Using video analysis with the Noldus system, we have conducted a study that brings together these three areas of questioning to go beyond establishing links between reading ability and signing ability. Using Noldus software, we analyzed the actual online behavior of deaf LSQ-signing children and their tutor.

Objectives of the study

The study examined how signers of Langue des signes québécoise (LSQ) [1] read a text in French.. We aimed 1) to go beyond the deaf reader's reading scores to find observable strategies for approaching print; and 2) to determine whether the use of fingerspelling, initialized signs and chaining during instruction influence reading outcomes, namely text comprehension.

Participants

Three case studies were conducted with three adolescent males who were profoundly deaf from birth. Their primary language was LSQ and their written mode of communication was French. Two tutors also participated in the study – the first was hearing with a native competence in French and a near-native level of signing competence; the second was deaf with native competence in LSQ and near native competence in written French. The choice of a hearing and a deaf tutor reflects our belief that signing behavior and teaching practices may differ as a result of language background as demonstrated by Padden and Ramsey [2,3].

Signing and reading measures

Participants' French reading was measured with 2 tests, the reading comprehension subtest of the French version of the Canadian Achievement Test, the *Test de rendement pour francophones* [4] and the French version of the Peabody Picture Vocabulary Test (PPVT, 1991) known as EVIP (*Échelle de vocabulaire en images Peabody*). A demographic questionnaire also yielded information regarding age of sign language acquisition, language of parents (deaf or hearing), language of siblings, use of LSQ, and reading behavior.

[1] *Langue des signes québécoise (LSQ) is the sign language used by deaf individuals in French parts of Canada.*

Procedure

Participants were asked to read and retell a text with the help of a tutor. They were then videotaped reading the story, signing "aloud" (with assistance if needed) and then retelling it (with prompts if needed). Stimulated recall sessions [5] were then conducted by a trained research assistant, with the assistance of an LSQ-French interpreter. Stimulated recall involves the teacher or student in an interview while watching the videotape of the reading and retelling activity he or she has participated in. The student and/or teacher is encouraged to comment on the reading activity process and to provide insight about how the student was able to achieve an understanding of the text with or without help. The interviews focused on what specific LSQ skills (i.e. fingerspelling, initialized signs, chaining) and what reading strategies (seeking meaning or word attack) were used by the teacher and student to understand the text. These sessions were videotaped for later analysis.

Each of the three participants interacted with both tutors on separate occasions reading a different text each time. This allowed a comparison of the strategies used by all three signers/readers in constructing meaning from text. An additional comparison was obtained from observing each of the three students interacting with both types of tutors. It was hypothesized 1) that the better reader would use a seeking meaning strategy for comprehending the text while the weaker reader would favor a word attack strategy; 2) that the deaf tutor would make more use of fingerspelling, initialized signs and chaining than the hearing tutor; and 3) that the weaker reader would achieve better understanding of the text when reading with the deaf tutor.

Analyses

All video footage was analyzed using the Noldus The Observer XT software. Both the story reading videos and the stimulated recall videos were converted into computerized files. The stimulated recall videos with the student and that of his teacher were then coded for strategies used in constructing the meaning of the text. Variables observed for each type of strategy include their frequency and their efficacy (did it work?) We then were able to compare strategies used by each teacher (deaf and hearing) with all three students. A second comparison between all three students sought differences between the better reader and the weaker reader in the study.

The use of video analyses and especially the Noldus The Observer XT software allowed us to gain a sense of each teacher's style of teaching through an in-depth analysis of the types and frequency of strategies used with all students, regardless of a student's profile. We were then able to show differences between the hearing teacher and the deaf teacher. For example, the deaf teacher relied more on *meaning-making* strategies while the hearing teacher used more *word attack* strategies. It was also possible to reanalyze the data with a focus on students leading to a finding that students used different strategies with each teacher, seemingly adapting to the style of the teacher.

Conclusion

Our presentation will describe the study in detail with a step by step look at the methodology used. Examples of videos and of analyses will be given as a way to illustrate how the Noldus The Observer XT software can be used effectively to analyze online data with a particular emphasis on sign language studies. Implications for future studies involving signing participants will be discussed.

References

1. Allen, T.E. & Schoem, S.R. (1997). *Educating deaf and hard-of-hearing youth: What works best?* Paper presented at the Combined Otolaryngological Spring Meetings of the American Academy of Otolaryngology, Scottsdale, AZ.
2. Padden, C. & Ramsey, C. (1998). Reading ability in signing deaf children. *Topics in Language Disorders*, **18** (4), 30-46.
3. Padden, C. & Ramsey, C. (2000). American Sign Language and reading ability in deaf children. C. Chamberlain, J.P. Morford and R.I. Mayberry (Eds.), *Language acquisition by eye*, (pp. 165-189). Mahwah, NJ: Lawrence Erlbaum Associates.
4. Sarrazin, G. (1996). *Test de rendement pour francophones*. Toronto: Psychological Corp.
5. Ericsson, K.A. & Simon, H. (1993). *Protocol analysis: Verbal reports as data*. Cambridge, MA: MIT Press.

Gesture coding with the NGCS - ELAN system

H. Lausberg[1] and H.Slöetjes[2]
[1]Institut für Psychosoziale Medizin und Psychotherapie, Friedrich-Schiller-Universität Jena, Jena, Germany, hedda.lausberg@med.uni-jena.de
[2]Max-Planck-Institute for Psycholinguistics, Nijmegen, The Netherlands, han.sloetjes@mpi.nl

The Neuropsychological Gesture Coding System (NGCS) is a tool for empirical gesture research that combines an objective kinetic and a functional analysis of gestures. The NGCS is tailored to the annotation tool ELAN.

The NGCS is based on neuroscience research. While previous gesture classification systems were developed on the basis of gesture observation and their categories were heuristical, the NGCS links the phenomenology to the neurobiology of gestures. The theoretical assumption behind the NGCS is that its main kinetic and functional gesture categories are differentially associated with specific cognitive (spatial cognition, language, praxis) and emotional functions. This implies that the different gesture categories are generated in different brain areas. Studies on split-brain patients in whom the production of specific gesture categories in the right and left hemispheres was examined [1-5] and functional Magnetic Resonance Imaging (fMRI) studies [6] provide evidence that, indeed, the different gesture categories of the NGCS are generated in different brain regions. A second theoretical assumption behind the NGCS is that gestures are medium of communication and expression that conveys information per se. Again, this hypothesis is strongly supported by recent split-brain research: in these subjects, a substantial amount of communicative gestures is generated in the separate right hemisphere, i.e., independently from left-hemispheric language production [1,5]. Thus, the NGCS coding is primarily performed without sound, and only after the gesture analysis has been accomplished, the relation to the verbal context can be analyzed. The NGCS gesture categories are defined by kinetic features only and not by interpretation of the verbal context.

The NGCS consists of two complementary modules: The kinetic analysis (Module I) comprises the coding of gesture laterality, structure, and location. Based on these criteria, Module I coding provides objective, kinetically defined gesture units. In Module II, the gesture units are examined concerning their function: pointing, spatial information, pictorial information, prosodic emphasis, emotional expression, and interactive regulation. The functional analysis has been developed on the basis of gesture types the existence of which different gesture researchers had agreed on [7-11]. In addition, the NGCS includes gesture types derived from apraxia research [12-14]. Each functional main category contains several gesture types. The twenty-three gesture types each are defined by a specific combination of kinetic aspects including hand shape, hand position, path, efforts, planes, kinesphere, gesture space, involvement of body parts other than hands, and gaze.

Thus, the NGCS enables to classify gestures based on their kinetic features alone, i.e., by video analysis without sound. With this procedure, the objectivity of the gesture types is high (Cohen's κ .87) [5]. As a behavioral and neuropsychologically grounded method, the NGCS can be used for a broad range of study designs such as spontaneous gesture production in everyday conversation, structured elicitation of gestures by stimuli, or examinations of patients with brain damage or mental disease. The NGCS is presented in form of an ELAN-template.

ELAN [15] is one of the annotation tools that have come into existence with the maturation of digital media. Originating from the early twenty-first century, it was intended to support audio as well as video annotation and to be versatile enough to accommodate different fields of research. Where some tools specialized in e.g. field linguistics or audio-only speech analysis, ELAN became a member of the group of multimodal annotation tools.

Written in the Java programming language, ELAN is available for Windows, MacOS and Linux. It complies with standards like Unicode and its file format is XML-Schema based. There is support for up to four video files that can either be integrated in the main window or detached in a separate, resizable window. In combination with milliseconds as the time-units, this allows for the best possible spatial as well as temporal inspection of the media.

A basic annotation is a textual label or tag attached to a segment, defined by a begin and an end time, of the media. In ELAN annotations are grouped on tiers or layers; a tier acts as a container for annotations that typically refer to the same kind of events. The user can define and create as many tiers as needed and tiers can be grouped hierarchically.

This is a necessity in most transcription tasks of today; be it in annotation of multiple modalities or in interlinearized annotation of speech. In the case of NGCS separate tiers are created for each gesture category that needs to be coded.

A few features of ELAN make it a particularly useful tool for the kind of research based on a solid theoretical foundation and a well defined classification system, like the NGCS with its more or less standardized gesture types. The first of these features is the Controlled Vocabulary facility. A controlled vocabulary is a user definable list of values that logically belong to the same category or the same group. By creating a number of controlled vocabularies, each holding the values of specific aspects of the gesture types, an incarnation of the theoretical model can be constructed. In the annotation editing process the vocabularies are used to present the proper list of values, thus accelerating the transcription task and making it less error prone. An ELAN Template file holds the skeleton of a transcription file: the definitions of the tiers, their types (i.e. sets of constraints [16]) and the controlled vocabularies. A template document can be used as the basis for new transcription documents. Documents based on the same template constitute a consistent and comparable set of resources. Sharing templates between teams highly improves interoperability of research resources.

The search system of ELAN provides the means to construct complex queries based on the structural and temporal relationship between annotations. The search results can be exported to a tab-delimited text file for further processing in a statistical analysis application.

Thus, the NGCS – ELAN tool enables researchers to conduct objective gesture coding and facilitates quantitative as well as qualitative gesture research.

References

1. Lausberg, H., Davis, M., Rothenhäusler, A. (2000). Hemispheric Specialization in Spontaneous Gesticulation in a Patient with Callosal Disconnection. *Neuropsychologia* **38**, 1654-1663.
2. Lausberg, H., Kita, S., Zaidel, E., Ptito, A. (2003). Split-brain patients neglect left personal space during right-handed gestures. *Neuropsychologia* **41**, 1317-1329.
3. Lausberg, H., Cruz, R.F., Kita, S., Zaidel, E., Ptito, A. (2003). Pantomime to visual presentation of objects: left hand dyspraxia in patients with complete callosotomy. *Brain* **126**, 343-360.
4. Lausberg, H., Cruz, R.F. (2004). Hemispheric specialisation in the imitation of hand positions and finger configurations. A controlled study in split-brain patients. *Neuropsychologia* **42**, 320-334.
5. Lausberg, H., Zaidel, E., Cruz, R.F., Ptito, A. (2007). Speech-independent production of communicative gestures: Evidence from patients with complete callosal disconnection. *Neuropsychologia* **45**, 3092-3104.
6. Lausberg, H., Heekeren, H., Kazzer, P., Wartenburger, I. Differential cortical mechanisms underlying pantomimed tool use and demonstrations with tool in hand. *Submitted.*
7. Efron, D. (1941). *Gesture and culture*. The Hague: Mouton.
8. Ekman, P., and Friesen W.V. (1969). The repertoire of non-verbal behavior. *Semiotica* **1**, 49-98.
9. Freedman, N., O'Hanlon, J., Oltman, P., & Witkin, H.A. (1972). The imprint of psychological differentiation on kinetic behaviour in varying communicative contexts. *Journal of Abnormal Psychology* **79**(3), 239-258.
10. Davis, M. (1991, rev. 1997). Guide to movement analysis methods*. Behavioral Measurement Database Services*, P.O. Box 110287, Pittsburgh, PA. USA.
11. McNeill, D. (1992). Hand and mind. What gestures reveal about thought. Chicago: *University of Chicago.*
12. Liepmann, H. (1908). Drei Aufsätze aus dem Apraxiegebiet. Berlin: *Karger*.
13. Goldenberg, G., Hagmann, S. (1998). Tool use and mechanical problem solving in apraxia. *Neuropsychologia* **36**(7), 581-589.
14. Goldenberg, G. (2001). Imitation and matching of hand and finger postures. *NeuroImage* **14**, 132-136.
15. http://www.lat-mpi.eu/tools/tools/elan
16. Brugman, H., Russel, A. (2004). Annotating Multi-media / Multi-modal resources with ELAN. *LREC.*

A speech adapted pattern recognition framework for measuring energetic states using low level descriptors and functionals

J. Krajewski[1], R.Wieland[1], D. Sommer[2], and M. Golz[2]

[1]*Work and Organizational Psychology, Univ. of Wuppertal, Wuppertal, Germany, krajewsk@uni-wupertal.de*

[2]*Neuro Computer Science, and Signal Processing, Univ. of Applied Sciences, Schmalkalden,Germany, m.golz@fh-sm.de*

Abstract

This paper describes a general framework for detecting affective and energetic states based on prosody, articulation and speech quality related speech characteristics. The advantages of this realtime approach are that obtaining speech data is non obstrusive, free from sensor application and calibration efforts. The main part of the feature computation is the combination of frame level based speech features (LLD) and high level contour descriptors (functionals) resulting in over 8,500 features per speech sample. In general the measurement process follows the speech adapted steps of pattern recognition: (a) recording speech, (b) preprocessing (segmenting speech units of interest), (c) feature computation (using perceptual and signal processing related features, as e.g. fundamental frequency, intensity, pause patterns, formants, cepstral coefficients), (d) dimensionality reduction (filter and wrapper based feature subset selection, (un-)supervised feature transformation), (e) classification (e.g. SVM, LDA, K-NN, MLP, HMM, GMM classifier; metaclassification e.g. bagging, boosting, voting, stacking), and (f) evaluation (e.g. 10-fold cross validation, leave-one-sample-out). The validity of this approach is briefly discussed by summarizing the empirical results of a sleep deprivation study.

Keywords

Computational Intelligence, Acoustic Features, Pattern Recogntion, Sleepiness Detection, Speech Emotion Recognition

Introduction

Many efforts have been reported in the literature for measuring biosignal based energetic states [1-3]. The real-time detection systems mainly focus on (a) oculomotoric data (eye blinking, eyelid movement, and saccade eye movement) [4], (b) EEG data [5] and (c) behavioral expression data (gross body movement, head movement, mannerism, and facial expression) [6] in order to characterize the energetic state. Apart from these promising advances in analysing facial and gestural expressivity, there has been recently renewed interest in vocal expression and speech analysis. Mainly this fact is promoted by the progress in speech science. Using voice communication as an indicator of sleepiness would have the following advantages: obtaining speech data is non-invasive, non obstrusive, free from sensor application and calibration efforts, “hands- and eyes-free”, and most important speech data is omnipresent in many daily life situations.

In this paper we describe a speech adapted pattern recognition framework in order to measure energetic states. Our attention is focused particularly on the processing step of feature computation. The rest of this paper is organized as follows: In Section 2 computing high level contour descriptor features is explained. The general speech adapted pattern recogntion framework is provided in Section 3, a brief summary of sleepiness detection results is given in Section 4.

High level contour descriptors as acoustic features

Frame level features (low-level descriptors)

Acoustic features can be divided according to auditive-perceptual concepts in prosody (pitch, intensity, rhythm, pause pattern, speech rate), articulation (slurred speech, reduction and elision phenomenons), and speech quality (breathy, tense, sharp, hoarse, modal voice) related features. Another distinction can be drawn from using signal processing categories as time, frequency or phase space domain features. Our approach prefers the fusion of perceptual features and purely signal processing and speech recognition based features without any known auditiv-perceptual pendants. Typical frame level based acoustic features used in emotion speech recognition and audio processing [7-9] are fundamental frequency (acoustic pendant to pitch; maximum of the autocorrelation function), intensity, duration of voiced/unvoiced segments, harmonics-to-noise ratio, position and bandwidth of 6 formants (resonance frequencies of the vocal tract depending strongly on its actual shape), 16 linear predictive coding coefficients, 12 mel frequency cepstrum coefficients (“spectrum of the spectrum”), and 12 linear frequency cepstrum coefficients (without the perceptual oriented transformation into the mel frequency scale).

Contour descriptors (functionals)

After splitting the speech signal into 10 ms frames and computing the above mentioned frame level features (Low-Level Descriptors, LLD; see [10]), the values of each frame level feature are connected to contours. This procedure results in about 57 speech feature contours (e.g. the fundamental frequency contour, the bandwidth of formant 4 contour etc.), which are joined by their first and second derivates (delta and delta-delta contours). Furthermore these 171 speech feature contours are described by elementary statisitics (linear moments, values and positions of extrema, quartiles, ranges, length of time periods beyond threshold values, regression coefficients, etc.), and spectral descriptors (spectral energy of low frequency bands vs. high frequency bands, etc.) resulting in about 8,500 high-level speech features (171 speech contours x 50 functionals).

Speech adapted pattern recognition framework

The measurement process follows the speech adapted steps of pattern recognition (see Table 1): (a) recording speech, (b) preprocessing (segmenting speech units of interest), (c) feature computation (using perceptual and signal processing related features, as e.g. fundamental frequency, intensity, pause patterns, formants, cepstral coefficients), (d) dimensionality reduction (filter and wrapper based feature subset selection, (un-)supervised feature transformation), (e) classification (e.g. SVM, LDA, K-NN, MLP, HMM, GMM classifier; metaclassification e.g. bagging, boosting, voting, stacking), and (f) evaluation (e.g. 10-fold cross validation, leave-one-sample-out).

Empirical validation results

We conducted a within-subject sleep deprivation design (N = 17; 8.00 p.m to 4.00 a.m). During the night of sleep deprivation a well proved, standardised self-report sleepiness measure, the Karolinska Sleepiness Scale (KSS) was used every hour just before the speech recordings. The verbal material consisted of a German phrase: “Ich suche die Friesenstraße“ [“I´m searching for the Friesen Street”]. The sentence was taken from simulated communication with a driver assistance system. The participants recorded other verbal material at the same session, but in this article we focus on the material described above. For training and classification purposes the records were further divided in two classes: sleepy (SS) and non sleepy (NSS) with the boundary value KSS ≥ 6. (46 samples NSS, 22 samples SS). During the night, the subjects were confined to the laboratory and supervised throughout the whole period. Between sessions, they remained in a room, watched DVD, and talked. Non caffeinated beverages and snacks were available ad libitum.

Table 1. *Processing steps and alternative specifications of the pattern recognition based speech acoustic measurement. The here used realizations of the pattern recognition specifications are printed in italics.*

Pattern recognition step	Specification
Recording	
Source of verbal material	*Human to human,* human to machine; *monolgue* vs. dialogue situation; speech databases (e.g. AEC, Sympafly, IFA, EMO-DB)
Speaking style	Vowel phonation, isolated words, connected speech, *read speech,* fluent speech, spontenous speech
Speech segment	Different vowels, different consonant type (fricative, stop, glide), consonant cluster, syllables, words, intonation unit, *phrases*
Recording situtation	Noisy vs. *noise subdued enviroment* (e.g. driving with open window vs. laboratory recording); rough vs. clean speech signal quality (e.g. telephon call, radio communication vs. clean recording in 22.05 kHz, 16 bit)
Preprocessing	
Segmentation	*Manual,* (semi-)automatical segmentation (e.g. MAUS system) of speech signal in units of interest
Framing	Size of frames (10-*20 ms*), degree of overlapping, window function (*hamming*, hanning)
Feature extraction	
Low level descriptors	*Fundamenal frequency, intensity-, harmonics-to-noise ratio, formant position and bandwidth (F1-F6), LPC, MFCC, LFCC, voiced speech segments, unvoiced speech segments*
Functionals	*linear moments, extrema values and positions, quartiles, ranges, length o time periods beyond threshold values, regression coefficients); spectral descriptors (spectral energy of low frequency bands vs. high frequency bands);* state space feature (largest lyapunov coefficient); automatic feature generation (genetic algorithms)
Normalization	*Individual speaker specific baseline correction*, age/ gender specific normalization; noise filtering
Dimensionality reduction	
Feature Subset Selection	*Filter based subset selection (correlationsfilter); wrapper-based subset selection* (forward selection, backward elimination, *genetic algorithm selection*)
Feature Transformation	*Unsupervised (principle component analysis)*; supervised (linear discriminant analysis)
Classification	
Classifier choice	Classification granularity (*binary* or multiple class prediction)*; 1-nearest neighbour, multi-layer perceptron, support vector machine,* linear discriminant analysis, hidden markov model, decision tree, gaussian mixture model; parameter optimization;
Metaclassification	Bagging, boosting, *voting*, stacking
Validation	
Evaluation strategy	*10-fold cross validation,* leave-one-speaker-out (LOSO), multiple-hold-ou

The averaged accuracy rates (ratio correctly classified samples through all samples) of three different classifiers (1-nearest neighbour, multi-layer perceptron, support vector machine) were over 80%. Due to the hypothezised sleepiness related physiological changes in cognitive speech planning, respiration, phonation, articulation, and radiation, the results for the reported classification performance were largely as could be expected. This is consistent with previous sleepiness related findings, that suggest an association of acoustic features [11,12] with sleepiness.

References

1. Kollias, S., Amir, N., Kim, J., Grandjean, D. (2004). *Description of potential exemplars: Signals and Signs of Emotion*. HUMAINE Human-Machine Interaction Network on Emotions.
2. Galley, N. (2007). On the way to a sleep warner. "Monitoring sleep and sleepiness with new sensors within medical and industrial applications", SENSATION IP 2nd International Conference, 4-5 June 2007, Chania, Greece.
3. Golz, M., Sommer, D., Mandic, D. (2006). Establishing a gold standard for drivers microsleep detection. *Proc. Int Conf Monitoring Sleep and Sleepiness (SENSATION 2006*, Basel, Switzerland), 29.

4. Caffier, P.P. (2002). The spontaneous eye-blink as sleepiness indicator in patients with obstructive sleep apnoea syndrome-a pilot study, *Sleep Medicine* **2**, 155-162.
5. Sommer, D., Chen, M., Golz, M., Trunschel, U., Mandic, D. (2005). Fusion of state space and frequency domain features for improved microsleep detection. W. Dutch et al. (Eds.): Int Conf Artifical Neural Networks (ICANN 2005), (pp. 753-759). Springer: Berlin.
6. Vöhringer-Kuhnt, T., Baumgarten, T. Karrer, K. & Briest, S. (2004). Wierwille's method of driver drowsiness evaluation revisited. *Proceeding of International Conference on Traffic & Transport Psychology.*
7. Schuller, B. (2006). *Automatische Emotionserkennung aus sprachlicher und manueller Interaktion.* [Automatic emotion recognition from verbal and manual interaction]. Dissertation, Technische Universität München.
8. Batliner, A., Steidl, S., Schuller, B, Seppi, D., Laskowski, K, Vogt, T., Devillers, L., Vidrascu, L., Amir, N., Kessous, L & Aharonson, V. (2006). Combining Efforts for Improving Automatic Classification of Emotional User States. In T. Erjavec & J.Z. Gros (Eds.), *Language Technologies, IS-LTC 2006*, (pp. 240-245). Ljubljana, Slovenia: Infornacijska Druzba.
9. Mierswa, I. , Morik, K. (2005). Automatic feature extraction for classifying audio data. Kluwe.
10. Vlasenko, B., Schuller, B., Wendemuth, A., Rigoll, G. (2007). Combining Frame and Turn-Level Information for Robust Recognition of Emotions within Speech. *Proceedings of Interspeech*, 2249-2252.
11. Nwe, T.L., Li, H**.,** Dong, M. (2006). Analysis and Detection of Speech under Sleep Deprivation. *Proceeding of Interspeech* 2006, 17-21.
12. Krajewski, J., Kröger, B. (2007). Using prosodic and spectral characteristics for sleepiness detection. *Proceedings of Interspeech,* 1841-1844.

Automatic detection of lameness in dairy cattle
Image features related to dairy cow lameness

Claudia Bahr[1], Toon Leroy[1], Xiangyu Song[1], Arno Pluk[1], Willem Maertens[2], Juergen Vangeyte[2], Annelies Van Nuffel, Bart Sonck[2], and Daniel Berckmans[1]

[1]*Division of Measure, Model & Manage Bioresponses (M3-BIORES), Katholieke Universiteit Leuven, Leuven, Belgium, e-mail, claudia.bahr@biw.kuleuven.be*

[2]*Institute for Agricultural and Fisheries Research Technology & Food Unit - Agricultural Engineering, Merelbeke, Belgium*

Lameness, an increasing animal welfare problem, has a negative impact on milk production, body condition and reproductive performance in dairy cows [1, 2 and 3]. The objective of this research is to develop and analyze image parameters correlated with expert gait scores that are applicable for continuous lameness detection. Experiments were done on ILVO farm in Ghent Belgium in August/September 2007. A camera recorded postures and movements of 10 lactating Holstein cows while passing the corridor to the pasture outside the barn. The gait of the cows was scored by experts on a scale of 1 to 3. Image features of track way overlap and spine arch were investigated by using labelled countour points. The trackway overlap can be described by the distance between the placement of the fore hoof on the floor and the placement of the hind hoof on the floor (see Figure 1). Within the spine arch 2 corresponding values can be calculated: the curvature of the back and the radius as a basis of the curvature (see Figure 2).

Results show that each calculated image feature has a relation to the expert scores on lameness. Regarding the parameter trackway overlap, The cows with walking problems have higher locomotion scores and greater positive trackway overlap values at least in one leg. The results of the parameter spine arch show that in between the 10 cows the curvature of the back becomes more convex with high gait score. The radius has a linear relation to the curvature and becomes smaller with an increasing convexity and a higher gait score.

It shows track way overlap and spine arch are useful image features for lameness detection. A further goal is the development of an automatic on-line lameness detection tool after analyzing more lame cows and more image features.

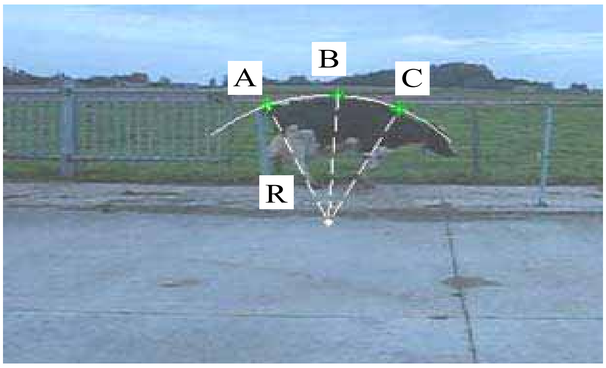


Figure 1. *Feature of spine arch curvature (κ) in the horizontal side view. κ = 1/R*

References.

1. Booth, C. J., L. D. Warnick, Y. T. Grohn, D. O. Maizon, C. L. Guard and D. Janssen. 2004. Effect of Lameness on Culling in Dairy Cows. *Journal Dairy Science* **87 (12)**, 4115-4122.
2. Holzhauer, M., H. Middelesch, C. Vartels and K. Frankena. 2004. Evaluation of a Dutch claw health scoring system in dairy cattle. *13th International Symposium and 5th Conference on Lameness in Ruminants, Maribor, Slovenija,* 18-21.
3. Vermunt, J., 2004. Herd lameness-a review, major causal factors, and guidelines for prevention and control. *13th International Symposium and 5th Conference on Lameness in Ruminants, Maribor, Slovenija,* 3-18.

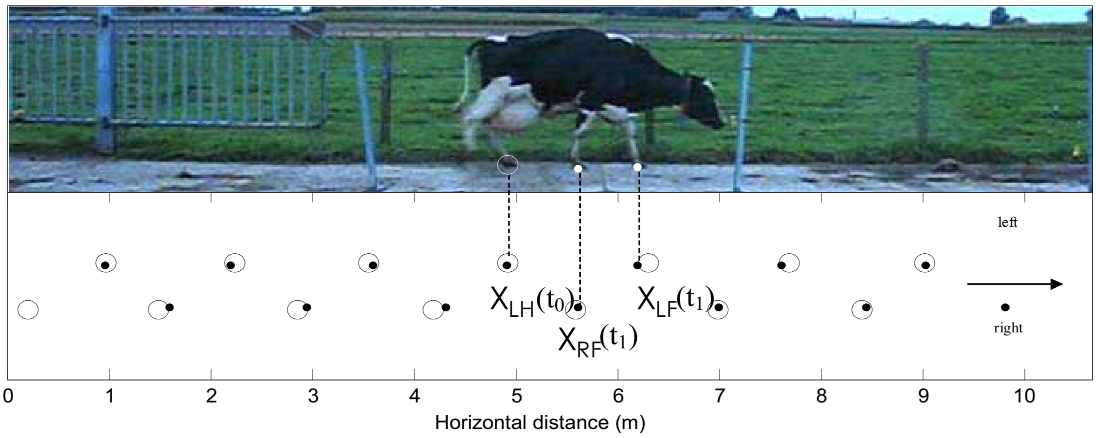


Figure 2. *Feature of track way overlap (Δ) in the horizontal side view.* $\Delta_L(t) = X_{LH}(t_1) - X_{LF}(t_0)$

An automated open field mouse gait analysis test

T. Leroy[1], S. Stroobants[2], J.-M. Aerts[1], R. D'Hooge[2], and D. Berckmans[1]
[1]*M3-BIORES, Katholieke Universiteit Leuven, Heverlee, Belgium, daniel.berckmans@biw.kuleuven.be*
[2]*Laboratory of Biological Psychology, Katholieke UniversiteitLeuven, Leuven, Belgium.*

Introduction

In the past, some efforts have been made to seek for automated systems to ease the job with monitoring laboratory animals. Commonly, these methods involve visually analyzing video recordings of the animal traversing a specially built runway [1,2,4]. Very recently, a more automated method, the "CatWalk" [3], uses a runway with internal reflection to increase contrast so image analysis software can recognize paw-like objects automatically. However, manual identification and labeling of the individual paws is still required, making the job time consuming. Also, all techniques are quite expensive because of the required material.

The objective of this research is to develop a system for fully automated mouse gait analysis using computer vision. The system automatically locates and identifies footprints of the mouse in camera images, acquired from below the transparent cage floor. From a sequence of footprints, the gait of the mouse is quantified using standardised gait parameters.

Materials & Methods

Data was acquired during 2 series of experiments on mice with a targeted disruption of the arsylsulphatase A (ASA) gene, with lowered coordinated locomotion abilities as a symptom. The first series were performed on 36 heterozygotes (12 females, 24 males) and 50 knock-out mice at the age of 6 months. A subgroup of 16 heterozygotes (9 females, 17 males) and 33 knock-out mice were used in a second series at the age of 12 months.

The mice were placed one at a time in the test setup, consisting of a plexi-glass cage (53x34.5x26cm) and two fluorescent bulbs for proper illumination (see Figure 1). The transparent cage allowed images to be captured from underneath the cage, so additional information could be obtained about the position of the limbs of the mice for gait reconstruction. Every mouse was recorded during 10 minutes.

Background subtraction and colour filtering were used to measure image features, such as the mouse's position, orientation, body outline and possible locations for the mouse's paws. A set of heuristic rules was used to prune implausible paw features and label the remaining ones as front/hind and left/right. After pruning implausible paw features, the paw features that are consistent over subsequent images are matched to footprints. Finally, from the measured footprint sequence, eight parameters were calculated to quantify the gait of the mouse: the stride lengths of the front & hind, left & right paws, the front & hind base and the left & right overlap.

Results

The system described above was used to automatically measure the gait of the 36 heterozygotes and 50 knock-out mice at 6 months. No gait information could be obtained for knock-out mice at the age of 12 months due to their decreased ability to move freely. Therefore, only results at the age of 6 months are regarded here. Furthermore, from the group at 6 months, 18 mice (5 heterozygotes and 13 knock-outs) were excluded from the analysis because a lack of movement.

There were clear distinctions between the gait pattern of heterozygotes and knock-outs (Figure 2). The maximal hind base (HB) displayed by knock-outs was significantly larger. The mean right overlap (RO) was significantly smaller for knock-outs than for heterozygotes. For the left overlap (LO), this difference also appeared but was less outspoken. Knock-outs also showed consistently larger mean and maximal stride lengths for left/right, front and hind paws (LFS, RFS, LHS and RHS).

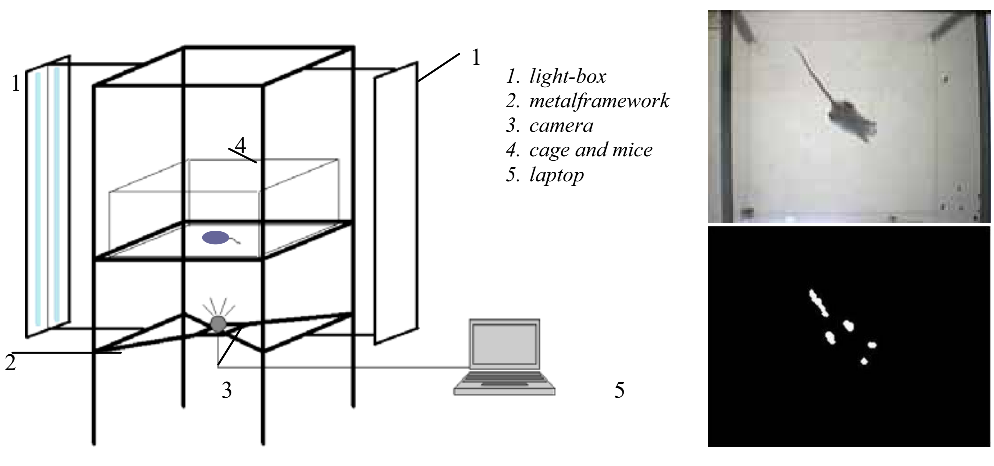


Figure 1. *A schematic overview of the experimental setup*

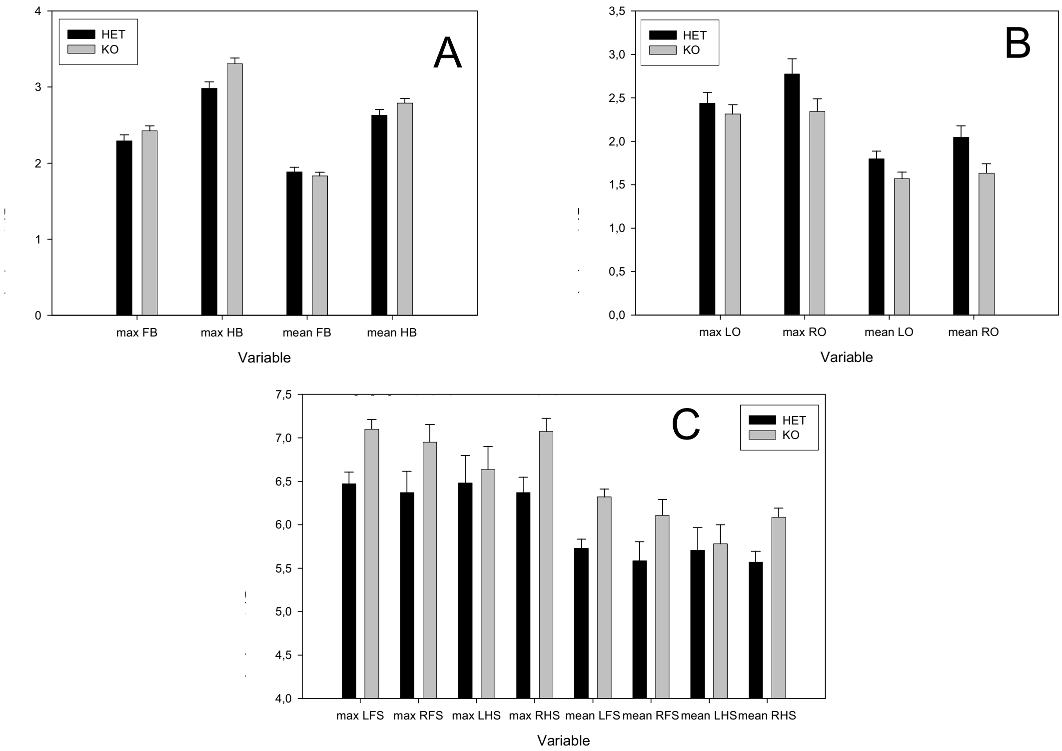


Figure 2. *Gait parameters of heterozygotes (HET) and knock-outs (KO) at the age of 6 months. Top left (A): The mean and maximal front base (FB) and hind base (HB). Top right (B): The mean and maximal left overlap (LO) and right overlap (RO). Bottom (C): The mean and maximal left front (LFS), right front (RFS), left hind (LHS) and right hind (RHS) stride lengths, respectively.*

References

1. Clarke, K. A., & Still, J. Development and consistency of gait in the mouse. (2001). *Physiology & Behavior*, **73**, 159–164.
2. Dijkstra, J. R., Meek, M. F., Robinson, P. H., & Gramsbergen, A. (2000). Methods to evaluate functional nerve recovery in adult rats: walking track analysis, video analysis and the withdrawal reflex. *Journal of Neuroscience Methods*, **96**, 89–96.
3. Hamers, F. P. T., Lankhorst, A. J., Van Laar, T. J., Veldhuis, W. B., & Gispen, W. H. (2001). Automated quantitative gait analysis during overground locomotion in the rat: its application to spinal cord contusion and transection injuries. *Journal of Neurotrauma*, **18**, 187–201.
4. Walker, J. L., Evans, J. M., Meade, P., Resig, P., & Sisken, B. F. (1994). Gait-stance duration as a measure of injury and recovery in the rat sciatic-nerve model. *Journal of Neuroscience Methods*, 52, 47–52.

Recognition of fatigue in gait patterns by means of support vector machines

D. Janssen[1], W. I. Schöllhorn[1], and K. Davids[2]

[1]*Training and Movement Science, University of Mainz, Mainz, Germany, djanssen@uni-mainz.de*

[2]*School of Human Movement Studies, Queensland University of Technology, Brisbane, Australia*

Introduction

Kinetic gait patterns seem to provide accurate information on individuality [1], pathologies [2], circadian rhythms [3] and emotions [4]. In former studies, recognition of fatigue in movement patterns was predominantly considered using linear classification methods with discrete movement parameters, for example in long distance running [5] or postural stability [6]. Effects on gait, induced by isokinetic exercises that fatigue local leg muscles have yet to be considered in research. Following previous work [7], the aim of this study was to train and test support vector machines on the ability of correctly classifying gait patterns before and after complete leg exhaustion. With this aim in mind the software package SVM-Light [9] was used. Support vector machines (SVMs [8]) overcome several problems that can be found in classical neural networks for example: dealing with small datasets and finding global minima. In addition to fatigue recognition, person recognition rates were calculated with all models.

Methods

Seven healthy participants (age: 25.4 ± 3.34 years) volunteered for the study. Kinetic data (vertical ground reaction forces) from a 1000Hz Kistler force plate (9821B, 60x40cm) were derived when participants walked barefoot for approximately 7 m at a self-determined walking speed registered by two pairs of double light-barriers. To establish baseline values, participants performed six gait cycles initially before they were required to completely exhaust their m. soleus and m. gastrocnemius with the aid of additional weights (45.7 ± 9.8 kg) for six trials, immediately before each treatment gait cycle. Data from twelve gait cycles per participant were recorded.

Vertical ground reaction forces were normalized by time (100 data points) and amplitude to remove or minimize the influence of speed and body weight on the recognition process. Thereby, amplitudes were normalized per participant and not per data sample. This method allowed retracing intra-individual changes and inter-individual comparisons at the same time, as weight information was removed from the data. Thereafter, two subsequent processing methods could be distinguished:

a) The time courses of all gait patterns were directly used as input for the SVMs.; b) From each gait pattern a synthetic model gait pattern was subtracted, that was constructed by calculating the mean of all twelve gait patterns of each individual participant. This was done in order to extract or reduce individual information from the gait data.

For the recognition of fatigue in gait patterns, the data samples were allocated to the two classes *fatigued* and *not fatigued* and presented to a SVM with a linear kernel. For person recognition a Multiclass-SVM with comparable parameters was used, as gait patterns had to be linked with the participants' names as classes for training purpose. Before the data were fed to the SVMs, a further normalization to the interval [0 to 1] was completed in order to prepare the data for the processing. In both cases cross validation was used for training and testing the SVMs. The classification procedure was performed 100 times with a training-testing ratio of 10:2 and a randomized allocation of data samples into training and test data. Recognition rates were calculated counting the misclassifications of the test data and expressing them as a percentage. Both recognition of fatigue and person recognition were accomplished using method a and method b.

Results

A fatigue recognition rate of 90.7% was achieved across all gait patterns when method *a* was chosen, whereas method *b* delivered a fatigue recognition rate of 95.2%. Considering the individuality of gait, a person recognition rate of 98.7% was achieved using method *a,* whereas method *b* delivered a rate of 41.7%.

There was no statistical difference between the participants' gait velocities for the *fatigued* and not *fatigued* conditions.

Discussion

With regard to recognition, individuality seems to have a stronger influence on gait pattern than fatigue effects. However, the suggested approach in method *b* offers the possibility of increasing the recognition of fatigued gait by up to 95.2%, when minimizing individual information. This approach allows for better inter-individual recognition of fatigue in gait patterns. Results have implications for detailed diagnosis of elite performance and training in sports and for therapeutic and rehabilitative interventions in clinical assessments.

References

1. Schöllhorn, W., Nigg, B., Stefanyshyn, D., & Liu, W. (2002). Identification of individual walking patterns using time discrete and time continuous data sets. *Gait & Posture* **15**, 180-186.
2. Kohle, M., Merkl, D., & Kastner, J. (1997). Clinical gait analysis by neural networks: issues and experiences. *Computer-Based Medical Systems, 1997. Proceedings., Tenth IEEE Symposium on*, 138-143.
3. Schöllhorn, W., Schaper, H., Kimmeskamp, S., & Milani, T. (2002). Inter- and intra-individual differentiation of dynamic foot pressure patterns by means of artificial neural nets. *Gait & Posture **16***, S172.
4. Janssen, D., Fölling, K., & Schöllhorn, W. (2006). Recognising emotions in biomechanical gait patterns with neural nets. *Journal of Biomechanics* **39**, S118.
5. Christina, K. A., White, S. C., & Gilchrist, L. A. (2001). Effect of localized muscle fatigue on vertical ground reaction forces and ankle joint motion during running. *Human Movement Science*, *20*, 257-276.
6. Salavati, M., Moghadam, M., Ebrahimi, I., & Arab, A. M. (2007). Changes in postural stability with fatigue of lower extremity frontal and sagittal plane movers. *Gait & Posture* **26**, 214-218.
7. Jäger, J., Alichmann, M., & Schöllhorn, W. (2003). Erkennung von Ermüdungszuständen anhand von Bodenreaktionskräften mittels neuronaler Netze. In *Biologische Systeme - Mechanische Eigenschaften und ihre Adaptation bei körperlicher Belastung* (S.. 179-183). Hamburg: Czwalina.
8. Vapnik, V. (1995). The Nature of Statistical Learning Theory. Berlin: Springer.

9. Joachims, T. (1999). Making large-Scale SVM Learning Practical. In B. Schölkopf, C. Burges & A. Smola (ed.). *Advances in Kernel Methods - Support Vector Learning.* MIT-Press.

Classical Gait Analysis reveals Regular Movement Cycles in Spontaneous Movements of Human Neonates

B. Aßmann, D. Loscher, and C. Niemitz
Institute for Human Biology, Free University Berlin, Berlin, Germany, birte.aszmann@gmx.de

Summary

We studied limb coordination in human neonates by applying the procedure of classical gait analysis – definition of gait cycles and normalization to percentages of these cycles - to kinematic data of spontaneous limb movements of human neonates in the supine position. Movement cycles of the limbs in the supine position were defined by maximal and minimal values of distance trajectories of the hands and feet from the centre of the torso, which is analogous to the definition of gait cycles of limb movements in the erect posture by initial contacts of the limbs with the ground. In addition to gait analysis, in which the gait cycles are commonly defined by the contact values of one limb, we extended the procedure to a parallel eightfold analysis by defining the minimal and maximal values of each of the four limbs as starting point for a new cycle, respectively. This new application of the classical analysis revealed regular movement cycles with integer ratios as well as symmetric cycles as known from adult gait patterns in the spontaneous movement behavior of human neonates in the first days of life.

Introduction

The investigation of spontaneous movements of neonates has experienced a huge transformation from traditional observation and categorization of e.g. primitive reflexes and general movements [1] to sophisticated methods from nonlinear dynamics, that e.g. describe spontaneous single-limb-activity by chaotic dynamics [2,3] and limit cycles [4]. In former studies, we analyzed the neonatal four-limb-system with methods from nonlinear dynamics, i.e. recurrence plot analysis and symbolic dynamics, and found evidence for processes of self-organization towards a system of transient reference configurations [5]. In search for the underlying mechanisms of the emergence of the documented system of four-limb-configurations, we adapted the classical gait analysis to analyse the neonate movements in the supine position and demonstrate how this traditional observation procedure can reveal hidden regularities in highly complex behavioral systems of several interacting and interdependent components.

Material and Methods

Subjects were six neonates, all healthy, normal, full-term infants, with Apgar scores of 8 or more at the age of 1 to 10 days (mean age = 3.2 days). Kinematic data collection used video recording with three synchronized cameras (50Hz) that focussed into a volume calibrated by a calibration frame. Kinematic data (4Hz) of movement relevant joints were computed using the Ariel Performance Analysis System (APAS) for time spans of 5- 20 minutes from each of the six infants on 2-3 different days. Distance trajectories of the endeffectors of the arms (wrists) and legs (ankles) from the centre of the trunk in the xz-dimension, and the minimal and maximal values of the latter, were calculated from the coordinate data. Movement cycles were defined from one maximum (or minimum) of the distance trajectory of one limb to the following maximum (or minimum) of the respective limb. The data for each cycle was normalized to percentage of cycle and the ratios to the maxima and minima values of the remaining limbs were calculated. The parallel procedure of maximal and minimal values for four limbs resulted in 8 cycle series, two for each limb respectively.

Results

With the parallel procedure of cycle analysis from maxima to maxima of each limb and minima to minima of each limb we got 8 cycle series. For each cycle of the eight series, the ratio of the second value of the respective limb (e.g. minima, if maxima defined the cycle) and the ratios of the remaining three limbs were calculated. Table 1 shows an example of 9 cycles defined by the maximal values of the distance trajectory of the left hand (LHmax). Cycle 22, 25 and 31 display integer ratios for all limbs involved. In cycle 22, the minimum of the left hand (LHmin) occured at 70% of the cycle, the maximum of the right foot (RFmax) and the right hand (RHmax) at 10% of the cycle as well as the minimum of the right hand (RHmin) at 100%, indicating synchronous movement of the latter with the left hand (LHmax). Simultaneous movements always occurred at 100% with the event defining the cycle (LHmax in table 1) as well as if concurrent values are displayed: e.g. the left hand minimum (LHmin) and the right foot minimum (RFmin) coincided in cycle 25 at 60% and the left hand minimum coincided with the left foot minimum in cycle 26 at 75%.

In general, regular and irregular cycles alternated within the whole movement episodes, with the percentage of regular cycles ranging from 21 to 34% for each cycle series. This means that altogether, ca. 80% of a movement episode was involved in regular cycles in reference to a minimal or maximal value of at least one of the limbs.

Table 1. *Illustrative section of a movement cycle analysis of 9 cycles of one movement episode (N=1898 cycles; Mean duration: 5s, SD: 5s) with maximal values of the left hand (LHmax) defining the onset and termination of a cycle. The left column 'Cycle' shows the number of the movement cycle. t displays the duration of the cycle from one maximum of the left hand (LHmax) to the following one. Percentages of the cycle at maximal and minimal values of the limbs are displayed: LH, left hand; RF, right foot; RH, right hand; LF, left foot; min, minimum; max, maximum.*

LHmax

Cycle	t [s]	LH min [%]	RF min [%]	RF max [%]	RH min [%]	RH max [%]	LF min [%]	LF max [%]
22	2.5	**70**		**10**	**100**	**10**		
23	9.25	29.72972973	8.108108108	32.43243243		35.13513514	16.21621622	37.83783784
			56.75675676	70.27027027	75.67567568	89.18918919	40.54054054	54.05405405
			78.37837838	91.89189189				
24	3	41.66666667	8.333333333	58.33333333	16.66666667			
25	6.25	**60**	**60**	**72**	**44**	**20**	**72**	
			80	**92**	**100**	**92**		
26	4	**75**	**31.25**	**43.75**		**18.75**	**75**	**31.25**
			81.25		**68.75**			
27	6.5	11.53846154	65.38461538	42.30769231	80.76923077	19.23076923	69.23076923	15.38461538
//								
31	3	**50**		**100**	**50**			
32	5.25	52.38095238	61.9047619		9.523809524	23.80952381		
					33.33333333	95.23809524		
33	2	**50**			**50**	**62.5**	**87.5**	**12.5**

Conclusions

Movement behavior often appears irregular and without any structure. In particular, if more than one channel of activity or information is involved. With our analysis, we want to show, how the parallel application of the classical procedure of gait analysis can unravel the clutter of single limb movements, general movements and reflexes by revealing that cycles with regular ratios as well as cycles with synchrony (100%) and 50% ratios (see cycle 31 and 33) - which are both common in adult gait patterns – were inherent in the movement flow and become obvious, if behavior is measured in respect to adequate reference points [5,7].

References

1. Prechtl, H.F.R., & Hopkins, B., (1986). Developmental transformations of spontaneous movements in early infancy, *Early Human Development* **14**, 233-238.
2. Taga, G., (2000). Nonlinear dynamics of the human motor control-real-time and anticipatory adaptation of locomotion and development of movements. *Proceedings of International Symposium of Adaptive Motion of Animals and Machines*, Montreal, Canada.
3. Taga, G., Takaya, R., Konishi, Y., (1999). Analysis of general movements of infants towars understanding of developmental principle for motor control, *Proceedings of IEEE SMC*, V678-683.
4. Thelen, E., Kelso, J.A. & Fogel, A. (1987b). Self-organizing Systems and Infant Motor Development. *Developmental Review* **7**, 39-65.
5. Aßmann, B., Thiel, M., Romano, M.C., & Niemitz, C., (2006). Recurrence plot analyses reveals a novel reference system in newborn spontaneous movements. *Behavior Research Methods,* 38 (3), 400-406.
6. Aßmann, B., Romano, M.C., Thiel, M., & Niemitz, C., (2007). Hierarchical organization in newborn spontaneous movements. *Infant Behavior and Development*, doi:10.1016/infbeh.2007.04.004.
7. Eilam, D., & Golani, I., (1989). Homebase behavior of rats (R. norvegicus) exploring a novel environment. *Behavioral Brain Research* **34**, 199-211.

The computation and effects of air traffic control message complexity and message length on pilot readback performance

O. Veronika Prinzo
Federal Aviation Administration, Civil Aerospace Medical Institute, Oklahoma City, OK, USA, roni.prinzo@faa.gov

Abstract

Ten years ago a comprehensive analysis quantified the types and frequency of air traffic control (ATC) communications. This submission briefly describes the computation of ATC message complexity, message length and their effects on pilot readback performance. Pilots experienced more difficulty reading back high complexity messages when on approach as compared to departure and the effects of message length were apparent only during approach.

Introduction

Humans are limited in the amount of information that they can effectively process, store, recognize, and recall. We gradually learn to organize sound into progressively larger groups by translating them into a verbal code [1]. As speech enters into verbal working memory (VWM) acoustically relevant sounds are extracted and encoded into phonemes that form syllables that are assembled to create words, phrases, clauses, and other constituents. These representations must be maintained in an active state (rehearsed) otherwise decay begins in about 2 s [2] or they are overwritten by new information. If the upper limit of VWM exceeds 5-7 chunks, problems may occur.

An utterance's complexity can be derived from the amount of information expressed in its constituents measured by the number of words, syntactic nodes, or phrasal nodes [3]. Utterances that exceed VWM capacity impose problems to listeners. The communication exchange between air traffic controllers and pilots is an excellent example.

In the US, air traffic controllers use a rigid set of words/phrases [4] to construct the messages they send to pilots who then read them back. During readback, controllers actively listen for accurate reproductions of the communication elements (CE) comprising their original messages. The presence of a mistake is a readback error (RBE).

ATC messages can contain multiple (CEs). Some words/ phrases serve as anchors that make a (CE) more precise in its interpretation. For example, "3-5-0" is ambiguous until it appears with an anchor — it can be interpreted as a heading, altitude, or speed. Thus degrees are associated with heading, knots with speed and descend/climb with altitude.

A complexity value (CV) is assigned to each anchor, numerical value, orientation (left, right, center), etc. according to the controller's phraseology usage. Furthermore, CV with larger values reflect the added complexity imposed by CEs with more information. To illustrate, 'three thousand five hundred,' 'one-zero thousand' and 'four thousand' most likely impose <u>quantitatively</u> different loads on VWM because 'three thousand five hundred' takes longer to pronounce and contains more words than 'four thousand' (e.g., articulatory loop [5]) and utilizes more capacity [1].

Methods

Audiotapes

Five US TRACON facilities provided a total of 28 hr 13 min 23 s of approach and 23 hr 56 min 32 s of departure communications.

Computing complexity

Each transmission was first parsed into CEs, labeled by speech act category and AT using the procedures developed by [6]. A CV was assigned to a) instructions/clearances speech acts that involved heading, heading modifier, altitude, altitude restriction, speed, approach/departure, frequency, route, and transponder ATs, b) advisory speech act that involved traffic and c) the altimeter portion of weather advisories.

Readback errors

A RBE is as an unsuccessful attempt by a pilot to repeat correctly the information transmitted by ATC. For example, ATC might transmit, "United Ten turn left heading two one zero." If the pilot read back either "three one zero" or "three six zero," it was coded as a substitution error since the numbers in the original heading included neither a three nor a six. If read back as "one two zero" it was coded as a transposition error since the correct numbers were spoken in an order different from the original. The absence of a number during readback was coded as an omission.

Results

Each readback was evaluated for accuracy and the number of errors recorded (e.g., a zero indicated no error while a value of 3 indicated 3 errors). There were 723 RBEs present in 688 pilot transmissions that were derived from 11,159 ATC transmissions. For the 6.2% faulty readbacks, 654 contained 1 error and another 34 contained 2 or more errors. Statistical significance was evaluated with $\alpha \leq .05$.

Message complexity

See Figure 1. Each ATC message was classified as either low ($\leq$ 09) or high ($\geq$10) complexity, paired with its readback, and mean RBE computed for each aircraft. A Sector (Approach, Departure) by Message Complexity (Low, High) ANOVA revealed that pilots produced more RBEs in an approach (Mean = .13) compared with a departure (Mean = .04) sector, [F(1,3700) = 129.00]. Also, more complex ATC messages had more RBEs (Mean = .17) than less complex messages (Mean = .04), [F1,3700) = 154.39]. There were more RBEs for approach high-complexity messages than departure high-complexity messages or low-complexity.

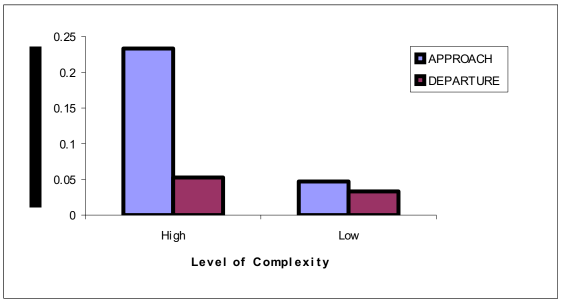


Figure 1. RBEs by sector and message complexity

Message length

See Figure 2. The results of the Sector (Approach, Departure) by Message Length (1AT, 2AT, 3AT, 4AT) ANOVA revealed more RBEs occurred when pilots were in the approach (Mean = .11), as compared with departure sectors (Mean = .03), [F(1,5599) = 78.48].

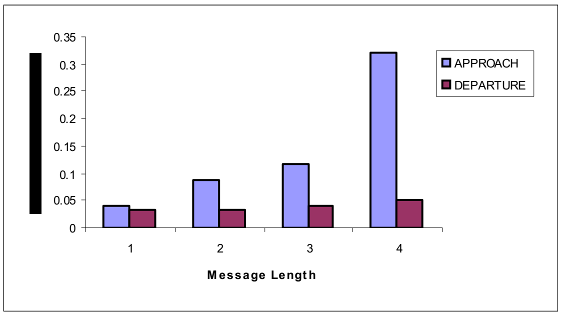


Figure 2. RBEs by sector and message length

The number of readback errors varied with the number of ATs, [F(3,5599) = 21.62]. The fewest readback errors occurred when ATC messages contained one AT (Mean = .04), no difference between messages with 2 or 3 ATs (2AT = .06; 3AT = .08), and messages with 4 ATs contained the most readback errors (Mean = .30). Figure 2 shows that as approach messages increased from one to between 2 and 3 ATs and 4 ATs the mean number of RBEs increased accordingly, [F(3,5599) = 21.62]. The effect of message length was apparent only for approach control.

Discussion

Whether by human or avionics, the accurate transmittal and receipt of information is necessary but not sufficient for communication to occur. Pilots and controllers, the human factors, must acknowledge or otherwise confirm that a common ground of understanding occurred between the source and its intended pilot recipient.

ATC message complexity and length can contribute to the vulnerability of pilot memory. The results provide evidence that RBEs may increase with increases in complexity and message length. Of particular interest was the finding that pilots experienced the most difficulty reading back ATC messages during the approach segment of their flight. Adding to their workload was the read back of a message with more than one AT or a complexity value ≥ 10 as evidenced by increased RBEs.

References

1. Miller, G.A. (1956). The magical number seven plus or minus two. Some limits on our capacity for processing information. *Psychological Review*, **63**, 81-97.
2. Baddeley, A.D., Thomson, N., and Buchanan, M. (1975). Word length and the structure of short-term memory. *Journal of Verbal Learning and Verbal Behavior*, **14***,* 575-589.
3. Wasow, T. (1997). Remarks on grammatical weight. *Language Variation and Change*, *9*, 81-105.
4. Federal Aviation Administration. (2004). FAA Order 7110.65P Air Traffic Control. www.faa.gov/ATPUBS/ATC/.
5. Baddeley, A.D. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences*, **4**, 417-23.
6. Prinzo, O., T. Britton, and A. Hendrix, 1995, *Development of a coding form for approach control/pilot voice communications,* Report no. DOT/FAA/AM-95/15, Washington, DC: FAA.

N = 1 effect sizes: Comparing models under autocorrelation

R. Manolov and A. Solanas
Department of Behavioral Sciences Methods, Faculty of Psychology, University of Barcelona, Barcelona, Spain, rrumenov13@ub.edu

Generalization of treatment effectiveness from single-case designs is possible whenever findings are replicated across units and integrated by means of meta-analysis. There has already been developed a diversity of measures summarizing the magnitude of the intervention effect: some designed for group studies (e.g., Cohen's *d*), while others specifically destined to N-of-1 designs (e.g., regression-based approaches like Gorsuch's trend analysis [1]; White, Rusch, Kazdin, & Hartmann's *d* [2], Allison & Gorman's model [3], and the non-regression Percent of Nonoverlapping Data; PND). The objective of the current study is to compare the abovementioned techniques in terms of effect detection under the presence of serial dependence and for different design lengths.

Method

The following design lengths were studied: a) N = 10; $n_A = n_B = 5$; b) N = 15; $n_A = 5$; $n_B = 10$; c) N = 15; $n_A = 7$; $n_B = 8$; d) N = 20; $n_A = 5$; $n_B = 15$; e) N = 20, $n_A = n_B = 10$; f) N = 30, $n_A = n_B = 15$.

In order to make possible the simulation of different data patterns (i.e., random fluctuation, level change, slope change, trend, and combination of effects), for each of the aforementioned design lengths data were generated according to the model presented in Huitema & McKean [4]:

$y_t = \beta_0 + \beta_1 * T_t + \beta_2 * D_t + \beta_3 * SC_t + \varepsilon_t$, where:

- y_t: the value of the dependent variable (behavior) at time t;
- β_0: intercept = 0.0;
- $\beta_1, \beta_2, \beta_3$: partial correlation coefficients;
- T_t: value of the time variable at time t (takes values from 1 to N);
- D_t: level change variable (equal to 0 for phase A and equal to 1 for phase B);
- SC_t: value of the slope change variable. $SC_t = [T_t - (n_A + 1)] * D_t$. The first n_A data points are equal to zero, while the following ones increment from 0 to ($n_B - 1$).
- ε_t: error term.
- The error term is generated according to: $\varepsilon_t = \varphi_1 * \varepsilon_{t-1} + u_t$, wherethe autoregressive parameter (φ_1) takes values from −0.9 to 0.9 with a step of 0.1.
- u_t follows a normal distribution with mean zero and unitary standard deviation

The values of the partial correlation coefficients were selected after several trials in such a way as to produce the same mean difference (equal to β_2) between phases for the shortest design ($n_A = n_B = 5$): $\beta_1 = 0.06, \beta_2 = 0.30, \beta_3 = 0.15$. Those values also allowed avoiding floor and ceiling effects.

For each of the experimental conditions defined by the combination of design length, autocorrelation level, and data pattern six effect sizes models were computed 100,000 times and were then averaged across all iterations. The effect size measurements were obtained in terms of R-squared, converting from *d* whenever necessary and using adjusted R-squared for Allison & Gorman's model as suggested by the authors. Solely the PND was not measured in the same scale and comparisons were made on the basis of visual inspection. Fortran 90 programs and NAGfl90 libraries' external subroutines were used for data generation (*nag_rand_seed_set* and *nag_rand_normal*) and multiple regression analysis (*nag_mult_lin_reg*).

Results

Gorsuch's model produced low effect size estimates, ranging from 0.01 to 0.06, concurring with Parker & Brossart [5]. This index produced R-squared that were affected by autocorrelation but did not differentiate between data patterns, and it was the model that showed poorest performance. Varying autocorrelation from −0.9 to 0.9 produced linearily increments in effect size estimates and the most affected techniques where the models of Allison & Gorman and White et al. The influence of serial dependence on PND had a U-shape and was less pronounced than in other techniques.

The regression-based techniques distinguished patterns in a lesser degree and only for long and balanced series. Moreover, the models of Allison & Gorman and White et al. produced seemingly too large R-squared. The two versions of Cohen's *d* performed better and a visual comparison reveals that PND differentiated the most between data patterns for the shortest designs.

Regards design length, as expected, longer data series led to a better differentiation between the effects present in the measurements, although designs such as $n_A = n_B = 5$ and $n_A = 5$, $n_B = 15$ were associated with greater R-squared than $n_A = n_B = 10$ or 15. Consistent with how data were simulated was the fact that changes in slope produced greater R-squared than changes in level.

Discussion

All of the models studied have been criticized on different basis: not being designed for N = 1 designs (Cohen's *d*), being too sensitive to outliers (PND), not taking into account slope change (Gorsuch, White et al.), producing too large R-squared (Allison & Gorman). The results of our simulation study show that simpler methods (like Cohen's *d* and PND) may be more effective than more sophisticated and conceptually more suitable (regression-based) methods. Further research is needed to explore optimal ways of simulating real behavioral data patterns, while other possible lines of future investigations include applying the effect size methods to designs controlling for extraneous variables (e.g., ABAB).

References

1. Gorsuch, R. L. (1983). Three methods for analyzing limited time-series (N of 1) data. *Behavioral Assessment,* **5**, 141-154.
2. White, D. M., Rusch, F. R., Kazdin, A. E., & Hartmann, D. P. (1989). Applications of meta-analysis in individual subject research. *Behavioral Assessment,* **11**, 281-296.
3. Allison, D. B., & Gorman, B. S. (1993). Calculating effect sizes for meta-analysis: The case of the single case. *Behaviour Research and Therapy,* **31**, 621-631.

4. Huitema, B. E., & McKean, J. W. (2007). An improved portmanteau test for autocorrelated errors in interrupted time-series regression models. *Behavior Research Methods,* **39**, 343-349.

5. Parker, R. I., & Brossart, D. F. (2003). Evaluating single-case research data: A comparison of seven statistical methods. *Behavior Therapy,* **34**, 189-211.

Measuring social reciprocity

A. Solanas[1], D. Leiva[1], V. Sierra[2], and Ll. Salafranca[1]
[1] *Department of Behavioural Sciences Methods, University of Barcelona, Barcelona, Spain*
[2] *Department of Quantitative Methods, ESADE-Ramon Llull University, Barcelona, Spain*
antonio.solanas@ub.edu

Social research has been mainly focussed on the individualistic approach, although it ignores the social context within individuals are embedded. This fact may explain why dyadic analysis has been increasingly applied to quantify group interactions. Regarding social reciprocity, the directionality of behavior in social interactions can be measured by the Directional consistency (DC) index [1]. This index only allows quantifying social reciprocity at group level, although social researchers are often interested in measuring social reciprocity at group, dyadic and individual levels. The Skew symmetry index can be used to measure social reciprocity at all levels [2]. A statistical test has been recently proposed to make decisions at group level for both statistics [3].

The Skew symmetry statistic is based on data at hand. That is, any sociomatrix is decomposed into its symmetrical and skew-symmetrical parts in order to quantify social reciprocity. Hence, the maximum value of the statistic depends on the number of recorded behaviors. The purpose of our study is to develop a modified version of the Skew symmetry statistic in which the maximum level of asymmetry could be obtained in any sociomatrix. This means that a standardized measurement should be derived.

Social reciprocity in groups can be represented by a matrix $\mathbf{\Pi}$, where the parameter π_{ij} denotes the probability of '*i* addresses a behavior to *j* in each interaction' Note that the parameters of the matrix $\mathbf{\Pi}$ allow defining a measurement of social reciprocity since they contain the main information to quantify dyadic reciprocity among all pairs of individuals. Hence, an index to measure global reciprocity in groups can be defined as

$$\Phi_r = \frac{tr\left(\mathbf{\Pi}'\mathbf{\Pi}\right) - \min\left(tr\left(\mathbf{\Pi}'\mathbf{\Pi}\right)\right)}{\max\left(tr\left(\mathbf{\Pi}'\mathbf{\Pi}\right)\right) - \min\left(tr\left(\mathbf{\Pi}'\mathbf{\Pi}\right)\right)}, \quad 0 \le \Phi_r \le 1$$

Social researchers do not know the value of Φ_r since they collect empirical data. Then, an estimator of asymmetry in social relations is required to obtain some information about social reciprocity. An estimator of the index Φ_r can be defined as follows:

$$\hat{\Phi}_r = \frac{\sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n} x_{ij}^2/c_{ij}^2 - \min\left(\sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n} x_{ij}^2/c_{ij}^2\right)}{\max\left(\sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n} x_{ij}^2/c_{ij}^2\right) - \min\left(\sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq 1}}^{n} x_{ij}^2/c_{ij}^2\right)} = \frac{\sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n} x_{ij}^2/c_{ij}^2 - m}{\max\left(\sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n} x_{ij}^2/c_{ij}^2\right) - m}, \quad 0 \le \hat{\Phi}_r \le 1$$

where x_{ij}, c_{ij} and n respectively denote the number of behaviors that individual *i* addresses to *j*, the amount of interactions in each dyad and the quantity of individuals in groups. It can be proven that the expected value and standard error of the statistic equal

$$E\left[\hat{\Phi}_r\right] = \frac{2\left(\sum_{i=1}^{n}\sum_{j=i+1}^{n} 1/c_{ij} + \sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n} \frac{\pi_{ij}^2\left(c_{ij}-1\right)}{c_{ij}} - m\right)}{n\left(n-1\right)-2m}$$

$$\sigma\left(\hat{\Phi}_r\right) = \frac{2\sqrt{\sum_{i=1}^{n}\sum_{j=i+1}^{n} \frac{q_{ij}+q_{ji}+2s_{ij}}{c_{ij}^4}}}{n\left(n-1\right)-2m}$$

$$q_{ij}+q_{ji}+2s_{ij} = 4c_{ij}\left(\pi_{ij} - 7\pi_{ij}^2 + 12\pi_{ij}^3 - 6\pi_{ij}^4\right) + 8c_{ij}^2\left(-\pi_{ij} + 6\pi_{ij}^2 - 10\pi_{ij}^3 + 5\pi_{ij}^4\right) + 4c_{ij}^3\left(\pi_{ij} - 5\pi_{ij}^2 + 8\pi_{ij}^3 - 4\pi_{ij}^4\right)$$

The expression for computing the $\hat{\Phi}_r$ statistic can be rewritten as

$$\hat{\Phi}_r = \frac{\sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n} x_{ij}^2/c_{ij}^2 - m}{n\left(n-1\right)/2 - m} = \sum_{i=1}^{n} \frac{\sum_{\substack{j=1\\ j\neq i}}^{n}\left(x_{ij}^2/c_{ij}^2 - \frac{1}{4}\right) - \frac{1}{2}\sum_{\substack{j=i+1\\ c_{ij}\ odd}}^{n}\frac{1}{c_{ij}^2}}{n\left(n-1\right)/2 - m} = \sum_{i=1}^{n}\hat{\phi}_i$$

Note that this expression enables to know each individual's contribution to asymmetry in social interactions. The expression for $\hat{\Phi}_r$ can also be written as follows:

$$\hat{\Phi}_r = \sum_{i=1}^{n} \hat{\phi}_i = \sum_{i=1}^{n} \sum_{j=1}^{n} \hat{\phi}_{ij} = \sum_{i=1}^{n} \sum_{j=i+1}^{n} \left(\hat{\phi}_{ij} + \hat{\phi}_{ji} \right) = \sum_{i=1}^{n} \sum_{j=i+1}^{n} \hat{\Phi}_{ij}$$

where $\hat{\Phi}_{ij}$ is a dyadic measurement.

To sum up, this study mainly deals with a new statistic for measuring social reciprocity at global level. Note that the expected value and standard error have been derived for this global measurement of social reciprocity. It allows social researchers to make suitable comparisons between empirical and expected statistic values under specific null hypotheses, often the null hypothesis corresponding to complete reciprocation. A Monte Carlo test is suggested to take statistical decisions regarding null hypotheses since the exact sampling distribution for the global statistic remains unknown. Additionally, the technique also enables social researchers to identify those dyads and individuals that contribute more to the lack of social reciprocity in groups.

References

1. van Hoof, J. A. R. A. M. & Wensing, J. A. B. (1987). Dominance and its behavioral measures in a captive wolf pack. In H. W. Frank (Ed.), *Man and wolf* (pp. 219-252). Dordrecht: Junk Publishers.
2. Solanas, A., Salafranca, Ll., Riba, C., Sierra, V., & Leiva, D. (2006). Quantifying social asymmetric structures. *Behavior Research Methods*, *38*, 390-399.
3. Leiva, D., Solanas, A., & Salafranca, Ll. (In press). Testing symmetry in social interactions: A comparison between the Directional consistency and Skew symmetry statistics. *Behavior Research Methods*.

An integrated Hidden Markov Model and Multilayer Perceptron classifier for improved behavior recognition in animals

Sobhita Meher[1] and Tucker Balch[2]
[1]*Department of Electrical Engineering, Institute of Technology, Banaras Hindu University, Varanasi, India, sobhita_meher@rediffmail.com*
[2]*College of Computing, Georgia Institute of Technology, Atlanta, USA, tucker.balch@gatech.edu*

Behavior in animals can in many cases be described as a sequence of actions shown. Such sequential data can be modeled with Hidden Markov Models [1]. Hence, Hidden Markov Models have been popularly used for behavior recognition [2] [4], including gesture recognition in human beings [4]. A Hidden Markov Model is a statistical model of a sequence of states that are not observable (called the hidden states) and that can be observed through another set of states that are produced by the underlying hidden sequence. Also, the unobservable process is a Markov process, that is, the current state is not effected by the path of past states save the most recent state. But, the lacuna with using Hidden Markov Models is that, each class has a model distinct from models of other classes, and the model of each class is trained separately to recognize only that class with the data set of that class only, so, Hidden Markov Models lack the property of interclass discrimination. Neural Networks are robust classifiers and can learn any kind of relationship between the input and the output data. And Multilayer Perceptrons [5] (a class of Neural Networks) are particularly suited to multi-class problems. And as they are trained both to recognize a particular class and not to recognize other classes, they are highly efficient in distinguishing between classes in a multi-class problem. But, Multilayer Perceptrons lack the capability to represent time-series data. So, given to the nature of the problem, and integrated classifier which has Hidden Markov Models in the first stage of the classifier and Multilayer Perceptrons in the second stage is best suited as it deals with the temporal nature of the data as well as the multi-class problem. We have used the classifier we have devised on simulated data for identifying the behavior of foraging ants. However, the same classifier can be used for behavior recognition in animals of any species, if we can identify the behavior as a sequence of actions performed by the animals and the actions are not directly observable but can be done through a set of observable features.

In our work, we have used the hand programmed environment of foraging ants inspired by the behavior of social insects, using the TeamBots simulation platform and motor schema based control developed by Guillory, Ngyuen, Balch, and Isbell [3]. Their experimentation generated data which we used for ant foraging behavior recognition. In the simulated environment there are 12 ants, 10 food objects, and 1 nest. We run the simulation till all the food objects were picked and all the ants in the simulation returned to the nest, which returned a sequence of hidden states for the 12 ants and the corresponding values of the set of four observable features.

Here, we can describe the behavior of ants as a sequence of different activities (the possible hidden states) they perform while searching food and bringing them to their nest. Activities typically shown are: loitering around the nest, exploring their environment, moving towards base, and moving towards target. Again, each of these states are associated with a set of 4 observable features, namely, something_in_gripper which gives information whether the ant is carrying food, at_nest which tells us about the location of the ant with respect to its nest, see_target which informs about the status of the visibility of the food (which is the target here) to the ant, and bumped which shows whether the ant bumped into the boundaries or another ant. The above mentioned activities in different sequences represent different behaviors shown by the ants. The four categories of behavior we have identified are: ant is loitering and not looking for food, when the ant looks for food and does not find any and finally returns to the nest, when the ant is at nest and not looking for food, when the ant finds food object.

The Hidden Markov Model initially predicts a behavior class and uncovers the possible sequence of states, and the second-stage classifier, the Multilayer Perceptron does the final classification. The Hidden Markov Model only recognizer yielded a recognition accuracy of 21.76%. And, the performance of the integrated classifier is 75.35%. From the results, it is evident that our classifier outperforms the traditionally used Hidden Markov Model classifier.

Since, in animal behavior recognition, behaviors can be described as sequences of actions and the actions are not directly observable but only through a set of observable features, our classifier can be used for behavior recognition in animals.

References

1. Rabiner, L. R. (1989). A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition. *Proceedings of IEEE* **77(2),** 257-286.
2. Feldman, A. and Balch, T. (2004). Representing Honey Bee Behavior for Recognition Using Human Trainable Models. *Adaptive Behavior* **12 (4)**, December 2004, 241-250.
3. Guillory, A., Balch, T., Isbell, C. (2006). Learning Executable Models of Behavior from Observations and Low Level Knowledge. *Proceedings of the Fifth International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS)(Hakodate, Japan, 2006)*, 795-797.
4. Brashear, H., Starner, T., Luckowicz, P., Junker, H. (2003). Using multiple sensors for mobile sign language recognition. *Proceedings of IEEE International Symposium on Wearable Computing (White Plains, WA, October 2003)*, 45-53.
5. Haykin, S. (1999). Neural Networks - A Comprehensive Foundation. New Jersey: Prentice-Hall Inc.

Observer agreement for timed-event sequential data: Time-based and event-based algorithms compared

Roger Bakeman[1], Vicenç Quera[2], and Augusto Gnisci[3]

[1]*Department of Psychology, Georgia State University, Atlanta, GA, USA, bakeman@gsu.edu*

[2]*Departamento de Metodología de las Ciencias del Comportamiento, Facultad de Psicología, Universidad de Barcelona, Barcelona, Spain, vquera@ub.edu*

[3]*Dipartimento di Psicologia, Seconda Università degli studi di Napoli, Caserta, Italy, augusto.gnisci@unina2.it.*

Among observational researchers, a single data logging approach is becoming increasingly standard [1]. Working with digital multimedia recordings displayed on computer monitors, observers depress keys to note onsets of events. Offsets may be explicitly logged as well, or inferred from the onset of a later coded event in the same mutually exclusive and exhaustive (ME&E) set. With such instrumentation, continuously alert observers (continuous sampling) log data in a way that allows frequency, duration, co-occurrence, and contingency information to be derived later.

The present report uses computer simulation to compare five algorithms for assessing observer agreement given *timed-event sequential data* (TSD) [2], that is, continuously-sampled, time-logged observational data of the sort just described. Two are time-unit based: time-unit kappa and time-unit kappa with tolerance; and three are event based: The Observer algorithm, the INTERACT algorithm, and the Generalized Sequential Querier (GSEQ) dynamic programming (DP) algorithm, respectively. The first and second are implemented in GSEQ; the first and third in The Observer Version 5.0 [3], and the first in Mangold International's INTERACT. The fourth will be implemented in future versions of INTERACT and the fifth in future versions of GSEQ. The GSEQ DP algorithm is an extension of a dynamic programming algorithm we developed previously for *event sequent data* (ESD; only sequence but no times recorded) [4].

All algorithms are based on an agreement matrix (or confusion) matrix. The matrix is by itself useful for observer training because it shows how observers agree and disagree; and although all algorithms use the known formula to compute kappa, none satisfy the assumption of independent tallies required by the classic Cohen's kappa [5]. Thus the kappas produced should not be confused with Cohen's.

Algorithms

Time-unit based algorithms tally successive time units; if the time unit is a second, the kappa table contains 300 tallies for a 300 s observation. Time-unit kappa with tolerance ($\kappa_{tolerance}$) tallies an agreement if a match is found in the other observer's record, not just for the same second but within a stated tolerance (time-window, often of 2 time units). Because values vary slightly, depending on which observer is considered first, its value is computed as the mean of two values. Event-based algorithms link events and add tallies (agreements or disagreements) to the kappa table based on which events are linked.

Depending on the algorithm, some events may be linked to more than one other event, some may be linked to a nil event (one observer records a code the other does not, an omission-commission error), or some events remain unlinked.

The Observer algorithm is based on an algorithm described by Haccou and Meelis [6], the INTERACT algorithm is a modification of The Observer one, and the GSEQ DP algorithm is based on the classic Needleman and Wunsch (NW) algorithm [7] for aligning sequences of nucleotides, with modifications proposed by Mannila and Ronkainen [8] and additional modifications by us. The NW algorithm belongs to a broad class of methods known as *dynamic programming*, which permit exact solutions without exhaustively exploring myriad possibilities. Users specify costs for exact agreements, specific disagreements, and omission-commission errors; depending on these costs, the algorithm then determines an optimal alignment between two sequences, a backward trace through dynamic programming matrixes defined by the algorithm identifies agreement, disagreement, and omission-commission errors.

Simulation

We developed a simulation program (OASTES, or Observer Agreement for Simulated Timed Event Sequences) that generates master records and then simulates how observers might code those records. The program lets us vary the number of codes (k), the variability of their probability and duration, and the observer accuracy, and then computes kappa for the five algorithms. Kappas, averaged over 1000 simulations, were computed for k = 5, 10, and 15; for low, medium, and high variability; and for 75%, 85%, and 95% observer accuracy. Results are shown in Figure 1. Averaged over the circumstances simulated, $\kappa_{tolerance}$ tended to be higher and GSEQ DP kappas lower, with The Observer and Interact kappas intermediate. Kappa with tolerance, compared to without, averaged .06 higher.

Conclusion

Of the two the time-based algorithms, we prefer $\kappa_{tolerance}$, not necessarily because it gives higher values, as expected, but because we think it reasonable not to count minor errors of timing on the order of just a few seconds. Moreover, eliminating such errors from the agreement matrix leaves those disagreements which are arguably more serious, and which can profitably serve as a basis for further observer training.

Of the three event-based algorithms, we think the GSEQ dynamic programming algorithm is more accurate. The Observer and INTERACT algorithms do not allow for omission-commission errors, The Observer may link even quite distant events, and INTERACT leaves some events unlinked. We think they overestimate kappa, thus it is not surprising that they both produce higher values than the GSEQ algorithm for the circumstances simulated. Moreover, the Needleman-Wunsch algorithm, on which the GSEQ algorithm is based, is conceptually sophisticated and has a firm basis in the literature.

Time-unit based kappas, with a tally for each time unit, likely overestimates how often observers are making decisions, whereas event-based kappas, with a tally for each agreement, disagreement, omission, and commission likely underestimates the number of decisions observers make. Sometimes (perhaps often) observers decide that an event is continuing and not changing to another event; such agreements are not counted by the event-based algorithms—

indeed, how often these private mental events occur may be unknowable. We conclude with a simple recommendation, not either-or but both. Report values for both a time-unit kappa and an event-based kappa; this range likely captures the "true" value of kappa. Similarly, provide observers with agreement matrixes for both a time-unit and an event-based kappa. Each provides somewhat different (time-based vs. event-based) but valuable information as to how observers are disagreeing, and so are useful in different ways as observers strive to improve their agreement.

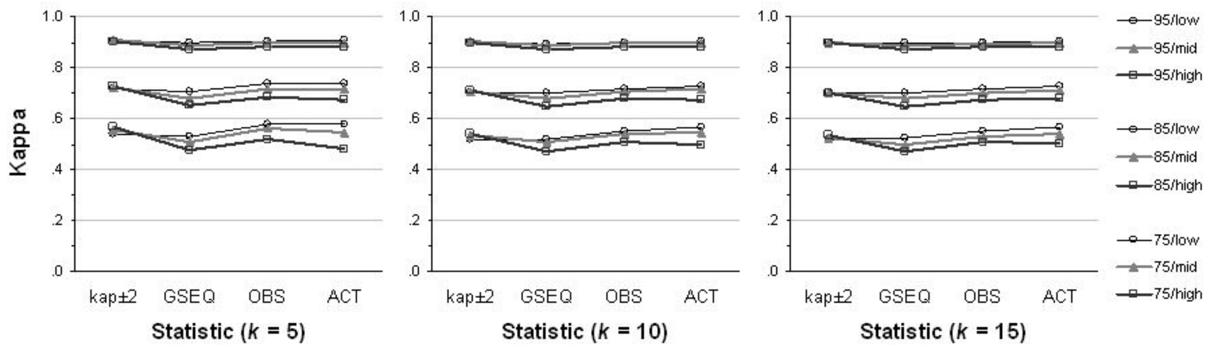


Figure 1. *Values for time-unit kappa (with 2 s tolerance), and as computed per the GSEQ dynamic programming, The Observer, and the INTERACT algorithms for k = 5, 10, and 15; observer accuracy = 75%, 85%, and 95%; and variability of code frequency and duration = low, moderate, and high.*

References

1. Jansen, R. G., Wiertz, L. F., Meyer, E. S., & Noldus, L. P. J. J. (2003). Reliability analysis of observational data: Problems, solutions, and software implementation. *Behavior Research Methods, Instruments, & Computers* **35,** 391–299.
2. Bakeman, R., & Quera, V. (1995). *Analyzing Interaction: Sequential Analysis with SDIS and GSEQ*. New York: Cambridge University Press.
3. Noldus Information Technology (2003). *The Observer: Professional system for collection, analysis, presentation and management of observational data. Reference Manual, Version 5.0*. Wageningen, The Netherlands: Author.
4. Quera, V., Bakeman, R., & Gnisci, A. (2007). Observer agreement for event sequences: Methods and software for sequence alignment and reliability estimates. *Behavior Research Methods* **39,** 39–49.
5. Cohen, J. A. (1960). A coefficient of agreement for nominal scales. Educational *and Psychological Measurement* **20,** 37–46.
6. Haccou, P., & Meelis, E. (1992). *Statistical analysis of behavioural data: An approach based on time-structured models.* Oxford: Oxford University Press.
7. Needleman, S. B., & Wunsch, C. D. (1970). A general method applicable to the search for similarities in the amino acid sequence of two proteins. *Journal of Molecular Biology* **48,** 443–453.
8. Mannila, H., & Ronkainen, P. (1997). Similarity of event sequences. In *Proceedings of the Fourth International Workshop on Temporal Representation and Reasoning. TIME'97* (Daytona Beach, Florida), 136-139

.

Effects of prevalence on observer agreement: a subjective assessment of working donkey behavior as an example

Charlotte C. Burn*[1], Joy C. Pritchard[1,2], and Helen R. Whay[1]
[1]University of Bristol, Department of Clinical Veterinary Science, Langford, Bristol BS40 5DU, UK
[2]The Brooke, Broadmead House, 21 Panton Street, London SW1Y 4DR, UK
*Corresponding author: charlotte.burn@bris.ac.uk

Kappa statistics are often used to assess the extent of observer agreement over subjective measures of behavior. They determine the degree to which agreement occurs compared with that expected by chance, so they are more conservative than raw percentage agreement scores. When used on measures with skewed prevalences, however, they become unreliable [1, 2]; with good agreement becoming almost impossible because the probability of observers agreeing purely by chance becomes so high. An alternative kappa calculation, 'PABAK', has been proposed to adjust for prevalence and observer bias [3], but this has been criticised for readjusting for the same factors that kappa is designed to control for [1]. There is no easy solution, so we suggest presenting prevalence indices and the raw percentage agreements alongside the kappa values, making kappa reliability more transparent [2, 4].

We have calculated the prevalence indices as the mean proportion of the most common classification relative to each alternative category, as described by a gold standard (in this case, the person who trained the observers). Thus, even for variables with many categories, the prevalence would be approximately 50% if categories were evenly distributed, but if the distribution was asymmetrical for any category, the prevalence index would increase. To aid interpretation of the prevalence indices we have divided them as follows: 50-59% = Well-balanced; 60-69% = Moderately balanced; 70-79% = Moderately skewed; 80-89% = Skewed; 90-100% = Highly skewed. It should be noted that these categories are only a guide, and their influence on agreement statistics will depend on the sample sizes used (even a slight skew could cause problems with small sample sizes).

We illustrate the above approach using the example of donkeys working in India. These animals have a high prevalence of welfare problems, and can appear unresponsive to the external environment and often demonstrate avoidance or aggressive behavior towards humans [5]. Five observers and their trainer (the gold standard) assessed the demeanour, lameness, and responses to humans of 80 donkeys.

The results are shown in Table 1. The percentage agreements for heat stress and gait were ≥ 98%, yet the overall ratings were Poor. This may mean that subjective assessment of these behaviors were indeed poor. Alternatively, however, the prevalence index showed that the gold standard scored 100% of the donkeys as showing no heat stress behavior and 98% as having abnormal gaits, so good observer agreement would have been almost impossible to detect. These measures can be contrasted against the response to observer approach; here the prevalence index was moderately balanced, so although the percentage agreements are much lower than for heat stress (≥ 74%), the overall agreement is Moderate. Demeanour showed Poor agreement, and this is likely to be a fairly reliable assessment of inter-observer performance because the prevalence index was only moderately skewed. For the other behaviors, agreement was moderate or substantial, despite skewed prevalences.

To conclude, kappa values are reliable when prevalence indices are moderately well balanced, and also when good agreement is obtained despite skewed prevalences. However, when prevalences are skewed, it remains unclear whether poor agreement ratings are due to the high probability of agreeing purely by chance, or due to genuinely poor agreement. This uncertainty should be acknowledged and, as is illustrated here, one approach is to provide prevalence indices alongside agreement ratings and percentage agreements.

References

1. Hoehler, F. K. (2000). Bias and prevalence effects on kappa viewed in terms of sensitivity and specificity. Journal of Clinical Epidemiology 53, 499-503.
2. Vach, W. (2005). The dependence of Cohen's kappa on the prevalence does not matter. Journal of Clinical Epidemiology 58, 655-661.
3. Byrt, T., Bishop, J., Carlin, J. B. (1993). Bias, prevalence and kappa. Journal of Clinical Epidemiology 46, 423-429.
4. Sim, J., Wright, C. C. (2005). The kappa statistic in reliability studies: use, interpretation, and sample size requirements. Physical Therapy 85, 257-268.
5. Pritchard, J. C., Lindberg, A. C., Main, D. C. J., Whay, H. R. (2005). Assessment of the welfare of working horses, mules and donkeys, using health and behaviour parameters. Preventive Veterinary Medicine 69, 265-283.
6. Landis, J. R., Koch, G. G. (1977). The measurement of observer agreement for categorical data. Biometrics 33, 159-174.

Table 1. *Inter- observer agreement results for a working donkey behavior assessment in India. Variables are arranged in ascending order according to the number of observers reaching criterion and their overall agreement ratings. Prevalence indices closer to 50% give more reliable kappa values. Kappa (k) values closer to 1.0 indicate better agreement, adjusting for that expected by chance. The rating scale is adapted from Landis and Koch [6] and Sim and Wright [4].*

Variable	Prevalence index (%)	Majority categories (%, if different to Prevalence index)	Minimum agreement (%)	Overall agreement (k)	Rating	Number of observers ≥ Moderate (total=5)
Heat stress (present / absent)	100	Absent	98	0	Poor	0
Gait (normal / abnormal)	98	Abnormal	98	0	Poor	0
Demeanour (alert / apathetic / depressed)	77	Apathetic (56%) or Alert (43%)	49	0.14	Poor	1
Response to observer walking down side (no interest / sign of interest)	73	Signs of interest	74	0.47	Moderate	3
Response to observer Approach (moves away / turns head away / no response / turns head towards / aggressive)	69	Turned head away (39%) or No response (38%)	65	0.58	Moderate	5
Chin contact (accepts/avoids	83	Accepted	86	0.67	Substantial	5
Tail tuck (no response to observer walking past rear / clamps tail down)	96	No tail-tuck	96	0.68	Substantial	5

Refined approach to register and analyze courtship behavior in *Drosophila*

N.G. Kamyshev, J.V. Bragina, N.G. Besedina, E.A. Kamysheva, and E.A. Timofeeva
Pavlov Institute of Physiology, St. Petersburg, Russia, nkamster@gmail.com

Introduction

The fruit fly, *Drosophila melanogaster*, is the model insect which has been widely and productively used in genetic studies since the beginning of 20th century. Starting from Seymour Benzer's initiative "from the gene to behavior", studies of significance of single genes for behavior and neural functions in the fruit fly have evolved into the powerful research field called "*Drosophila* neurogenetics" [1, 2]. The sequencing of *Drosophila* genome revealed, among others, the fact that there is no established function for about 60% of genes predicted by sequencing [3]. Many of these genes may appear to be "behavioral", i.e. their function may be ascertained by examination of behavior in genetically modified flies. Courtship is the most complex behavior in *Drosophila*. It involves mutual exchange of multiple sensory signals between male and female, as well as complex motor programs [4]. Being so rich in engaging various physiological (and, correspondingly, molecular) mechanisms, courtship is ideal to reveal the maximum of potential behavioral deviations in genetically modified flies using only one test system. This, however, requires refined approach to registration and detailed analysis of courtship behavior. Here, we present the computer program, which makes possible such a detailed analysis, and give an example of its usage. The program is destined to register and analyze the conjugate male and female ethograms using previously made video recordings of courting pairs.

Program description

The video file in MPEG-I format is viewed and analyzed using 6 modes. Mode 1 is used just to watch video. Seven speeds are available: from 1 to 25 frames per second. Watching may start from any particular frame selected with the track bar. One could also look through the clip frame by frame, either in forward or backward direction, or jump 10 frames forward or backward. This functionality is preserved for all other modes. Mode 2 is for male ethogram registration. Here the experimenter has an opportunity to fix the exact moments of beginning of male behaviors in male ethogram by pressing respective keys. The keyboard layout is constantly viewed. While the clip proceeds, pressing specific keys is immediately reflected in the male legend. If female ethogram has been already registered, the changes in female behavior are also reflected in the female legend. If one has pressed significant keys during clip demonstration, the program asks whether this information should be saved in male ethogram. The ethogram is saved as the sequence of behavior codes with corresponding frame numbers. The experimenter could perform repeated registrations starting from any frame and finishing at any frame, the matching piece of the ethogram will be automatically rewritten after confirmation. Mode 3 is for female ethogram registration. Mode 4 is for ethograms revision. While clip proceeds in this mode, the registered changes in both male and female behavior are reflected by corresponding legends. One could check the accuracy and accordance of male and female ethograms. Corresponding changes in the ethograms may be made in modes 5 and 6, which are used for frame by frame editing. Additional features of the viewing and registration interface include searching a certain frame by its number and searching the successive appearances of a particular behavioral element in the ethograms.

The ethograms are automatically saved in the same directory and under the same name as the video file, either with "mal" or "fem" extension. They are used by various analytical tools, each of which has its own interface. In all tools after selection of a group of video files to be analyzed, the program will use the corresponding ethograms, either male or female, or both, when needed. If the ethogram for a given video file is absent, the file is automatically excluded from the analysis, and the user is notified. Creating data matrix for an experimental group is the most general tool used for data analysis. Six matrices are created simultaneously: 3 for males and 3 for females. The first and second matrices present percentage of time occupied by various behavioral elements in each individual and frequency of their initiation (per 100 s). The third matrix lists durations of separate episodes of the same behaviors. The behaviors registered and analyzed for males include: orientation, pursuit, wing vibration either coupled or uncoupled with locomotion, tapping, licking, attempted copulation, locomotion and rest unrelated to courtship, preening with either forelegs or hindlegs directed towards a female, nondirectional preening, kicking (as attack), kicking (as the attack repulse), a leg put aside, backward wing flapping, wing "scissoring". The female behaviors, in addition to general to both sexes elements, include: escape (runaway) from a courting male, rest in neutral or "face-to-face" position, extruding the ovipositor coupled or uncoupled with locomotion. The matrices also include various combinations of the above behaviors (courtship index, *etc.*). The data matrices may be copied from the text editing control to datasheet of another program, such as Microsoft Excel, for further analysis and creating graphs. Other analytical tools allow to perform statistical comparisons of all behavioral parameters between independent and dependent samples and to analyze male behavior in its dependence on female's actions, and *vice versa*.

Example of usage

As an example, the analysis of behavior of the wild type and P-insertional mutant $Ent2^{P124}$ is presented, showing dependence of duration of the attempted copulation and male courtship intensity after attempted copulation upon ovipositor extrusion by a female during the attempted copulation.

Acknowledgements

The study was supported by federal and regional grants from Russian Academy of Sciences and from Russian foundation for basic researches.

References

1. Sokolowski, M.B. (2001). *Drosophila*: genetics meets behaviour. *Nat Rev Genet.* **2**, 879-890.
2. Vosshall, L.B. (2007). Into the mind of a fly. *Nature* **450**, 193-197.
3. Adams M.D., *et al.* (2000). The genome sequence of *Drosophila melanogaster*. *Science* **287**, 2185–2195.
4. Greenspan, R.J., Ferveur J.F. (2000). Courtship in *Drosophila*. *Annu Rev Genet.* **34**, 205-232.

Three statistical methods to analyze breed differences in the behavioral response to a challenging situation in kittens

P. Marchei[1], S. Diverio[2], N. Falocci[3], J. Fatjó[1], J.L. Ruiz-de-la-Torre[1], and X. Manteca[1]
[1]*Departament de Ciència Animal i dels Aliments, Facultat de Veterinaria, Universitat Autonoma de Barcelona, Bellaterra, Spain, paola.marchei@uab.es*
[2]*Dipartimento di Scienze Biopatologiche ed Igiene delle Produzioni Animali e Alimentari, Facoltà di Medicina Veterinaria Perugia, Italy*
[3]*Servizio Studi, Consiglio Regionale dell'Umbria, Perugia Italy*

Introduction

In a previous study [1], the behavior of Oriental, Siamese and Abyssinian kittens (OSA-k) was compared with that of Norvegian Forest kittens (NFO-k). Interesting breed differences emerged. In the present study, further statistical analsyses were carried out at different levels of complexity to supply a broad set of results. Outcomes were matched toghether in order to satisfactorily explain behavioral differences observed in the kittens during the response to a challenging situation.

Materials and methods

Pure breed kittens were supplied by ANFI (Associazione Nazionale Felina Italiana) and SCC (Serenissima Cat Club) cat breeders. 43 OSA-k and 39 NFO-k were exposed weekly to a 12 minute Open Field Test (OFT) from the 4^{th} to the 10^{th} week of age. The OFT arena was a white rounded box (1,80 m diameter); on the floor there was a cylindrical container as novel object (NO). In each session a randomly chosen kitten was introduced in the arena. After six minutes (First Part of the test - FP), a potentially aversive stimulus (AS) was produced by the sudden opening of the NO and the abrupt exit of a metal spring. The test lasted 6 more minutes (Second part of the test - SP) and the kitten was then removed from the arena.

Data collection procedures

<u>Animal identification</u>: schedule, pedigree name, mother, sex and coat colour were noted for each animal.

<u>Behavioral data</u>: all OFT were recorded with a video camera (Sony HandyCam®) placed outside the arena. Videos were downloaded on a PC using Microsoft® Windows® MovieMaker 5.1. Two trained operators registered the duration of behaviors over a total of 720 seconds (360 seconds - FP; 360 seconds - SP), according to predefined categories by *focal animal sampling* [2].

Data analysis

All data were analyzed using SAS 9.1 (SAS 2003, USA). The following statistical analyses were carried out:

Principal Components Analysis (PCA): data were introduced as mean values across animals to explore association between behaviors, breed, week and FP/SP.

Analysis of Variance with repeated measurements *(ANOVA)*; to compare mean durations of each behavior among breeds, FP/SP and weeks (considered as a factor) and their interactions;

Generalized Random Effect Model, with two different specifications for the error term:

- *Gamma distribution,* with durations as dependent variable, in order to evaluate behavoiural trends along weeks, defined as a covariate; quadratic term was included to assess the existence of non-linear trends;
- *Binomial distribution,* where the dependent variable was a *dummy* variable describing the presence or absence of a single behavior; the model assessed changes in the probability of performing a certain behavoiur in breeds, FP/SP and their interactions.

Results and discussion

The *PCA* showed two components explaining over 90% of the variability in behavioral patterns. The first principal component could describe "anxiety", with "anxious" behaviors like *crouched* and *escape attempts* on one extreme and "relaxed" behaviors, like *walking and exploring floor* on the other; anxious behaviors appeared to be associated with SP in both breeds, suggesting that AS was actually frightening; NFO-k were associated with *escape attempts*, OSA-k with the *crouched* posture. The second principal component seemed to be linked to "interest in exploration" or "curiosity", as there was a regular increase of explorative behaviors through the axis. Behaviors indicative of curiosity were on the same side of the plot of NFO-k during FP and SP.

The hypothesis was that OSA-k would passively face a challenging situation. In NFO-k, both novelty and fear would induce an active strategy. Our suggestions were validated by the further statistical analysis as described above.

ANOVA showed that, during FP, NFO-k explored more than OSA-k (p<0.001), but exploration decreased only in NFO-k after AV (p<0.0001); during first contact with NO, NFO-k spent more time exploring it (p=0.021) than OSA-k; but they increased the latency to enter in contact with it after AS (p=0.011). After AS, NFO-k presented a higher increase in resting postures and spent more time in escape attempts (p=0.034). In general, OSA-k receded more than NFO-k (p<0.0001), but in the latter breed the behavior increased after AV (p=0.003). The *Binomial Model* showed that in both breeds, for all behaviors and among weeks, the longer time animals spent performing a behavior (results from ANOVA), the higher the number of animals that exhibited it.

Analyzing evolution of behaviors with the *Gamma Model,* interesting changes appeared with time: in NFO-k emerged a strong variation in exploration levels, that increased during the first weeks, achieved its maximum value during central weeks and then decreased in the last weeks. In OSA-k such a modulation did not appear clearly, denoting less interest for the environment. Conversely, in OSA-k this pattern appeared for the behavior *escape attempts,* whereas in NFO-k this behavior was maintained at the same higher level until the last week.

Conclusions

Globally examined, our results match well with similar results on temperament obtained in other species [3]. A future research development would be to use a *Multivariate Regression Model* to analyze behavior as a whole, exploring the covariance structure among several behaviors during a fixed time period and the extent of correlation between this structure and the other explanatory variables.

References

1. Marchei, P., Diverio, S., Fatjó, J., Ruiz-de-la-Torre J.L., Manteca, X. (2007). Breed differences in behavioural development in kittens. *Proceedings of the 6th International Veterinary Behaviour Meeting (Italia, 17-20 June)*.
2. Martin, P., Bateson, P. (1993). Measuring animal behaviour. An introductory guide. Cambridge, UK: Cambridge University Press.
3. Kazlauckas V., Schuh J., Dall'Igna O.P., Pereira G.S., Bonan C.D., Lara D.R. (2005). Behavioural and cognitive profile of mice with high and low exploratory phenotypes. *Behavioural Brain Research* **162**, 272-278.

Induction of a social differentiation in human groups submitted to an experimental situation based on the rodent diving-for-food model

H. Schroeder[1], A.M. Toniolo[2], and D. Desor[1]

[1]Neurosciences Comportementales, URAFPA, INRA UC340, Nancy-Université, Vandœuvre-les-Nancy, France, henri.schroeder@scbiol.uhp-nancy.fr

[2]Department of Psychology, Nancy Université, Nancy, France

Introduction

In nature, many forms of organization have been observed in animal societies in response to an environmental constraint. Many mammal species adapt to the constraint by developing a complex and flexible system of social behaviors. This requires the implication of a kind of cognition named the social cognition that allows to everyone to construct mental representations of other members of the social group, as well as of the relations between them. Knowledge of this web of relations is essential for the expression and adaptation of individual behavior. In rodents, the diving-for-food model is an experimental situation that is suitable to provide a large set of information for studying the implication of this type of cognitive processes in the social organization of groups faced to an environmental constraint. This is a model in which the food accessibility is made difficult by progressively immersing the only way of access with water, leading to the emergence of a behavioral differentiation in groups of rats faced with this constraint [1, 2, 3]. To be more precise, a single food source is located at the end of an aquarium, so that rats have to dive and swim under water to get the food, and they have to go back to the cage where congeners are present, to eat it. In these conditions, a behavioral differentiation based on social interactions appears between Carrier rats which dive and bring the food back to the home cage, and the Non-Carrier animals which never dive and obtain their food by stealing it from the other members of the group. The present communication will describe the experimental situation, the procedure and the behavioral characteristics of a human model that consists of a transposition of the rodent diving-for-food model in every respect.

Material and methods

The transposition of the diving-for-food model in humans is therefore directed towards a playful aspect. Thus, the aim of this experimental situation is the hoarding of points by each participant within a group of 6 players. The points consist of small light plastic crosses (2 g) issued in bags of forty units. They can be obtained in two ways: either by overcoming a motor skill task, or by "stealing" them to their partners after a confrontation through a strategy game. The final appropriation of the points is effective when crosses have been introduced into the slot lid of a plastic box, each player having its own box. The experiment presented as a game seems to be an element of motivation enough for everyone because all participants have made the different tasks proposed to get points.

Experimental situation

The experience takes place in a 20 m^2 room within a delimited area divided into 3 different areas: the area of the motor skill task, the area of the strategy game, and the area of individual boxes. The experiment is presented as a game which reveals the purpose, the rules and the different elements of the game.

The only way for introduction of the points into the game is to win to the individual motor skill task. Each participant have to wade through a wire spiral (1.20 m long composed of 9 spires of 4 cm in diameter) with a metal ring and to avoid contact with the wire. A bell reports any failure. In the event of success the subject receives a bag of 40 pieces, i.e., 40 points.

The second solution to get points is to challenge one of a subject who holds a packet of points through a strategy game named “Power Play 4”, and if succeed in, to obtain the bag with the points it contains from the holder. This game requires cognitive abilities to develop strategies more or less effective and can compensate for a failure to the motor skill task.

Procedure

The experimental study of one group of 6 subjects includes 4 sessions of 30 minutes each, interspersed with periods of 15 minutes of rest. During the first 4 sessions, the number of spires to wade progressively increased from 1 at the first session to 9 at the last one in order to improve the difficulty to get points. The 5th session is a replica of session 4: it is designed to verify the stability of the social organization of the group.

Behavioral scoring

As for rats in the diving-for-food model, behavioral items related to the hoarding of points are scored. A first set of variables assesses the performance of each subject (number of points in the box at the end of the session, total number of hoarding points). A second set of items is related to the appropriation of the points (time of point possession at the box, number of periods of possession ending with the exhaustion of the points). The third one corresponds to the way of getting points (number of attempts and number of success to the motor skill task, number of point thefts they committed, number of challenges for stealing points). Finally, the fourth set of items is related to the loss of points (number of point thefts they experienced, number of challenges for protecting points). Behavioral data are analyzed by means of principal component analysis in order to identify the statistical dimensions around which behaviors are organized. Then, individuals are classified into groups of similar behavioral profiles by means of a hierarchical agglomerative cluster analysis of Euclidean distances using Ward’s method.

Results and perspectives

The results obtained from all the groups tested until now [4] show the emergence of a behavioral differentiation in 3 different profiles as in the rat diving-for-food model. The Non-Carrier subjects never succeeded in the motor skill task and have to get their points by stealing it to the holders. The two other profiles correspond to the Carriers. They present some abilities for getting points through the psychomotor task and can be separated into two subtypes based on their efficacy in protected their points in the strategy game. The more efficient participants are designed as the Autonomous and the less ones as the Suppliers. This model in humans could represent new perspectives to better understand the interactions between individual and social factors that lead a subject to adapt to his social environment.

References

1. Colin, C.; Desor, D. (1986). Behavioural differentiation in groups of rats subjected to a difficulty of access to the food. *Behavioural Processes*, **13**, 85-100.
2. Krafft, B., Colin, C.; Peignot, P. (1994). Diving-for-food: a new model to assess social roles in a group of laboratory rats. *Ethology*, **96**, 11-23.
3. Schroeder, H., Desor, D. (2005). The behavioural differentiation between the Carrier and the Non-Carrier profiles in groups subjected to the diving-for-food situation: a complex social model to study anxiety in rodents. *Proceedings of Measuring Behavior 2005, 5th International Conference on Methods and Technics in Behavioural Research (Wageningen, August 30-September 02, 2005), 390-393.*
4. Toniolo, A.M. (2005) Résolution de problème individuel en situation collective : une dynamique complexe à l'épreuve de l'expérimentation. *Revue Européenne de Psychologie Appliquée,* **55**, 29-42.

In PC and XBOX war games, are military personnel’s fixations patterns different compared to civilians’?

T. Falkmer[1,2], J Khalid[1,4], M R Afroz[1,4], H Hasewinkel[3], A Sjörs[1], and Joakim Dahlman[1]

[1]Rehabilitation Medicine, Department of Neuroscience and Locomotion, IKE, Faculty of Health Sciences, Linköping University, Sweden, torbjorn.falkmer@hhj.hj.se, anna.sjors@inr.liu.se, joada@inr.liu.se

[2]School of Health Sciences, HHJ, Jönköping University, Sweden, torbjorn.falkmer@hhj.hj.se

[3]The Swedish Rescue Services Agency (SRSA), Gothenburg, Sweden, hasewinkel@gmail.com

[4]Department of Information Technology, Uppsala University, Sweden, jkhalid@gmail.com, rayees80@gmail.com

Background

The use of digital tools for training and learning purposes has been growing in recent years, mainly due to a rapid technology development in the gaming and entertainment area [1]. Advances in game artificial intelligence, better human/computer interface, and stunning graphics capabilities have contributed to make video games interactive, adaptive and intelligent learning tools [2]. Gaming has been used effectively in areas where learning or training is difficult and costly in real and practical life [3]. However, appropriateness, practicability and effectiveness should be considered before using a particular game in a particular area of training [4].

For a long time, Armed Forces around the world have been using digital tools for training purposes [1]. These tools include at one end high cost, full-sized, sophisticated simulators with hydraulics, wall-sized video screens and realistic cockpits. At the other end, they include low-cost, relatively easy to develop and deploy computer based video games [5]. Many video games nowadays are specifically developed for military training purposes [5]. Such tools may be developed particularly for training and recruiting purposes, as is the case with the First to Fight-game.

Visual search strategies are critical elements for successful performance in the games [1]. Hence, it is important to study, before deploying any digital training tool, how war fighters search for visual information in a war game. The aim of the study was to compare military trained personnel with civilians, with respect to fixation duration and number of fixations on objects of interest (OoI) and areas of interest (AoI), and whether or not the gaming console had an impact on the fixations patterns.

Methods

In total, 20 military and 20 civilian subjects played ‘First to Fight’ on two different consoles, i.e. XBOX and PC, with a head mounted eye tracking device. All subjects played the game twice, i.e. on both consoles, in a fully balanced order. A questionnaire was used to get their views on the game environment. In total, 27,081 fixations were generated through a centroid mode algorithm [6] and analyzed; 13,101 fixations (48%) from Military personnel and 13,980 (52%) fixations from Civilians. The analyses were made manually, frame-by-frame. χ^2-test was used for dichotomized variables, paired samples and independent samples t-test for comparison of normally distributed interval variables. The α-level was set to .05 for all tests.

Results

The results showed that military trained personnel’s visual search strategies were different form the civilians’ when it comes to video game based war. Fixation durations were, however, equally short, i.e. about 170 msec, for both groups. To our surprise the military trained personnel’s fixation patterns were less orientated towards objects of interest and areas of interest than the civilians’; as shown in table 1 (χ^2 =82.35, df 1, p<0.001) and (χ^2 =42.60, df 1, p<0.001) respectively, the underlying mechanisms remaining unclear.

Military training was apparently not advantageous with respect to playing the ‘First to Fight’ video game. The PC console seemed to be advantageous over XBOX, with respect to fixation duration (t=4.595, df 27,079, p<0.001), and Military trained personnel fixated more often on OoI when they play PC first than when they played XBOX first (χ^2 =57.10, df 1, p<0.001), whereas no difference was found with respect to AoI and the order of consoles played. Furthermore, their fixations were shorter when they played PC first than when they played XBOX first (paired sample t=-198.68, df 13,100, p<0.001). Civilians, on the other hand, showed a contradictory pattern. They fixated more often both on OoI (χ^2 =5.93, df 1, p<0.05) and AoI (χ^2 =40.94, df 1, p<0.001) when the played XBOX first. However, their fixations were shorter when they played PC first than when they played XBOX first. (paired sample t=-198.16, df 13,979, p<0.001).

***Table1**. The distribution of the 27,081 fixations with respect to OoI/OO and AoI/OA among military personnel and civilians, and with respect to the two platforms.*

	Fixations on OoI	Fixations on OO*	Total	Fixations on AoI	Fixations on OA**	Total
XBOX	5,215, 35.6%	9,449, 64.4%	14,664 100%	10,537, 71.9%	4,127, 28.1%	14,664 100%
PC	4,620, 37.2%	7,797, 62.8%	12,417, 100%	8,605, 69.3%	3,812, 30.7%	12,417, 100%
Total	9,835, 36.3%	17,246, 63.7%	27,081 100%	19,142, 70.7%	7,939, 29.3%	27,081, 100%

** Other Objects, **Other Areas*

Conclusions

Military trained personnel's visual search strategies are different from civilians' when it comes to video game based war. Their fixation patterns were less orientated towards objects of interest and areas of interest than the civilians'. The PC console was advantageous over XBOX.

References

1. Khalid, J., Rayees Afroz, M. (2007). Visual Search Patterns. *Master Thesis, Uppsala University, Uppsala*
2. Bisson, C., Luckner, J. (1996). Fun in learning: The pedagogical role of fun in adventure education. *Journal of Experimental Education,* **19(2)**, 112.
3. Allen, R. W., Rosenthal, T. H., Hogue, J. R., Anderson, F. G., Pelz, C. A. (1999). Simulation for research, design review, assessment and training. *Paper presented at the TRB 3D Transportation, Orlando, FL*.
4. Falkmer, T. (2005). Truck and bus driver training, can simulation contribute? In L. Dorn (Ed.), Driver Behaviour and Training Volume II (II ed., Vol. II, pp. 93-104).
5. Hasewinkel, H. (2006). A Blueprint for Using Commercial Games off the Shelf in Defence Training, Education and Research Simulations. Licentiate *Thesis No 1283 Linköping University, Linköping.*
6. Falkmer, T., Dahlman, J., Dukic, T., Bjällmark, A., Larsson M. (2008). Fixation identification in centroid versus start-point modes using eye tracking data. *Perceptual Motor Skills*, **106,** 710-724.

Me, my SIMS, and my real me: Implicit measure of attitude using The SIMS

Sia Tjundjing
Faculty of Psychology, University of Surabaya, Surabaya, Indonesia, std@ubaya.ac.id

Background

Since its appearance in 2000, the coherence of the game play in a computer simulation game, The SIMS with real life behavior of its players (so called Simmers) has been discussed widely [2]. As recently as 2007, Griebel [2] published his empirical research, which supports the possibility to apply The SIMS to portray an individual's psychological characteristics and value system through the Big Five personality profile and Swartz Value Survey (SVS). Though recognized as a breakthrough in behavior research and measurement, Griebel still based his research on self-report, either to measure personality characteristics or play pattern in The SIMS.

This research is keen to step further by adding a new measurement method, which focusses on the results of the game play observation [3]. The main difference is the addition of the new measurement method, i.e. not only self-report but self-report and observation. In this research, the duration and latency of an activity are observed and recorded to reflect a more comprehensive gameplay profile.

Hypothesis

Hypothesis 1: Psychological characteristics, such as Big Five personality profile, individual value system, and academic performance, are correlated to gameplay.

Hypothesis 2: Participants psychological profile would be projected in their gameplay. These projections are manifested through their duration in doing certain activity, such as finishing assignments, sleeping, eating

Methods

Participants

Participants were 33 psychology students from University of Surabaya. Participants were recruited while following a program for students who registered taking the final term paper.

Materials

Personality factors were measured using Big Five Scale-Indonesian version [4]. Individual value system was measured using Portrait of Value Questionnaire/PVQ [5]. Academic Performances were measured by students GPA and class performances. The SIMS 2 gameplay was measured using The SIMS 2 University Survey. This 34 items questionnaire is a modification of Griebel's 25-items The SIMS 2 Survey. The SIMS 2 survey asked participants to report detailed information about how they played *The SIMS 2.*

Procedure

Initial session

Participants were administered a set of background information survey followed by informed consent, before they completed the BFSI and PVQ. This session lasted about half an hour.

The SIMS session

Tutorial session. To ensure they have had an adequate understanding to controlling their Sims, all participants should complete a tutorial session.

Create a student-session. At this session, they have to create their own Sim. Participants choose a name, physical, and psychological characteristics for their sim.

Briefing session. Before playing, participants receive brief information regarding life as a student in The SIMS.

Play session. Participants were told to choose their room and major and then begin to play *The SIMS 2* for 2 final exam periods. It's about 6 SIMS' days (approximately 3 real time hours).

The SIMS 2 University Survey Session. After completing the second final exam, the participants were administered *The SIMS 2 University* Survey, which took approximately 10 to 15 minutes to complete.

Data analysis

To complete the measurement, participant's gameplay were recorded into a bitmap picture (BMP) format every three seconds resulting about 4000 pictures for each participant (software SnagIt, version 8). Those pictures were scanned using self-made picture recognizing software to generate a database consisting pictures details, on-going activity, and sim's daily needs fluctuation.

Data analyses for the SIMS gameplay were conducted through two stages: all listed activities and focusing only on "do assignment" activity. All listed activities were analyzed merely based on total duration (7 SIMS' day) of each activity. "Do assignment" activity were analyzed using two indicators namely duration (daily and total) and latency (at what time the activity begins each day). The database were combined with the questionnaires' database to be processed further in the statistical data analyses (bivariate correlation tests).

Results

Hypothesis 1. The BFSI, PVQ, as well as the academic performance factors score reveal positive correlations (.301 until .493) with The SIMS' University Survey factors score. The PVQ factors score also shows negative correlations (-.307 until -.313) with The SIMS' University Survey factors score.

Hypothesis 2. The BFSI, PVQ, as well as the academic performance factors score reveal positive (.310 until .676) and negative (-.311 until -.939) correlations with the SIMS' gameplay.

Table 1. *Comparison with the previous study (Griebel, 2007)*

Aspects	Griebel (2007)	Current
Play duration	30 SIMS' days (30 real time hours in a six-week span)	6 SIMS' days (3-4 real time hours)
Participants	30	33
Tutorial session	Facultative	Compulsory
Number of sims played	1-4 sims	1 sim
Version	The SIMS 2	The SIMS 2: University expansion pack
Scenario	None	College life of a student who live in a comfortable dormitory
Location	Free	13-rooms dormitory (Sim State University)
Data source	Questionnaires Open ended questions	Questionnaires Observation of the The SIMS gameplay
Weaknesses	Demand characteristics Social desirability Sample size & representativeness	Same as Griebel's weaknessess Language barrier Mild correlations at the questionnaire sections
Strenghts	High correlations at the questionnaire sections	High correlation at the observation sections

References

1. Becker, D. (2004). *Controlling 'The Sims'*. Retrieved March 7, 2007, from http://www.news.com/2008-1043_3-5366681.html
2. Griebel, T. (2007). Self-Portrayal in a Simulated Life: projecting Personality and Values in The Sims 2. *Gamestudies* **6(1)**. Retrieved March 10, 2007, from http://gamestudies.org/0601/articles/griebel
3. Wittenbrink, B., & Schwarz, N. (Eds.). (2007). Implicit measures of attitudes. New York: Guilford Press.
4. John, O. P., Donahue, E. M., & Kentle, R. L. (1991). The Big Five Inventory--Versions 4a and 54. Berkeley: University of California, Berkeley, Institute of Personality and Social Research
5. Schwartz , S. H. , Melech , G. , Lehmann , A. , Burgess , S. , Harris , M. , & Owens , V. (2001). Extending the cross-cultural validity of the theory of basic human values with a different method of measurement . *Journal of Cross-Cultural Psychology*, **32**, 519-542 .

Measuring media exposure: Age and memory complications

Brian G. Southwell

School of Journalism and Mass Communication, University of Minnesota, Minneapolis, MN, USA, south026@umn.edu

Mass media campaign evaluations tend to depend on self-reported exposure measures in assessing effects. Many of these measures rely upon participant recognition of campaign material. Rather than accepting the utility of such measures at face value, however, we should probe their limits. We can predict at least one theoretically important limit on the basis of what we know about age-related memory changes.

Here we can make important distinctions between memory for specific detail and memory for the general gist of content. Age appears to bear a different relationship to detail memory, which corresponds to recognition-based tasks, than it does to gist memory, which we might expect to correspond to open-ended, free recall tasks. Relationships between information exposure, engagement, and memory vary with age [1]. Age predicts recognition error, for example [2, 3]. At the same time, free recall does not suffer the same decline as recognition memory and it also appears that older adults tend to depend on gist-based memory more than their younger adult peers [4, 5].

All of these patterns have implications for mass media exposure measures employing self-reported memory tasks. Specifically, the utility of a recognition item as an indicator of past exposure should decline as audience members age, especially among the most elderly, whereas recall items should not witness the same decline. A prominent evaluation measure, in other words, might suffer from differential validity across age groups.

Methods

Analysis of data from a science communication project evaluation offers support for these measurement hypotheses. Experimental data gathered as part of a project evaluation boasts two important strengths: participants were recruited by random digit dialing (from the Buffalo, NY, USA, 10-county Designated Market Area) and then randomly assigned to a level of science news story exposure. This means we have data from a reasonably generalizable sample (at least of the designated market area in question) and yet also can assess the impact of relatively carefully controlled exposure. We know who ostensibly had an opportunity to see material and can use that information to validate self-reported exposure measures that rely on memory tasks.

For this study, we focus on a subgroup (n = 347) from the total study sample (n = 667) who answered both recognition and open-ended, free recall questions a week after having the opportunity to view science news content. Randomly assigned participants entered either one of two treatment conditions (offered exposure to a partial or full diet of science news stories embedded in a week's worth of news programming) or a control group (offered the same general programming without any of the science stories in question). Assigned exposure consequently ranged from zero to six to 14 stories.

After receiving programming on a DVD or VHS tape, respondents were asked by telephone for their recognition of past engagement with a selection of the news stories in question roughly a week after they had been asked to view the shows. All respondents, regardless of condition, were asked about six different stories. We included in the recognition task a balance of stories, with three from those shown to the partial and full treatment conditions and three from those shown only to the full treatment condition. That permitted a recognition index which ranged from zero (stories recognized) to six (stories recognized). Before answering recognition items, participants also initially described the media content they had watched. A team of coders specifically looked for reference to one of the six stories in question, allowing us to create a similar free recall index.

Results

Results support our hypotheses. In short, the relationship between experimentally assigned physical exposure and subsequent self-reported recognition waned among adults 70 and older relative to their younger counterparts whereas the relationship between physical exposure and self-reported free recall for specific content remained essentially the same for both groups. Among those under 70 years old, the correlation between story recognition and physically assigned exposure was .67, p < .01, n = 303. Among those 70 and older, however, the correlation between story recognition and physically assigned exposure was substantially lower, r = .32, p < .05, n = 44. The picture was quite different for free recall. The correlation between story recall and physically assigned exposure was similar in both groups (under 70 r = .35, p < .01, n = 303 and 70 and older r = .40, p < .01, n = 44).

Conclusions

Evidence reported here supports the general hypothesis that recognition measures for mass media efforts do not work the same for all age groups. These results are centrally relevant to social science evaluation of media-based efforts involving older audiences. In short, they suggest that we might not be accounting for exposure in a uniform way across groups, a possibility that suggests either measurement noise or unanticipated theoretical complication in terms of potential effects. Accordingly, these results should give pause to practitioners and evaluators designing research to investigate media effects in elderly groups and warrant further investigation into the role that age plays in media exposure assessment.

References

1. Nussbaum, J. F., Pecchioni, L. L., Robinson, J. D., Thompson, T. L. (2000). *Communication and Aging* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.
2. Brainerd, C. J., Reyna, V. F. (2005). *The Science of False Memory*. New York: Oxford University Press.
3. Law, S., Hawkins, S. A., Craik, F. I. M. (1998). Repetition-induced belief in the elderly: Rehabilitating age-related memory deficits. *Journal of Consumer Research* **25**, 91-107.
4. Schneider, W. (2000). Research on memory development: Historical trends and current themes. *International Journal of Behavioral Development* **24**, 407-420.
5. Craik, F.I.M. (2006). Brain-behavior relations across the lifespan: A commentary. *Neuroscience and Biobehavioral Reviews* **30**, 885-892.

The Effects of apparatus design and test procedure on learning and memory performance of C57BL/6J mice on the Barnes maze

Timothy P. O'Leary and Richard E. Brown
Psychology Department, Dalhousie University, Halifax, NS, Canada, tpoleary@dal.ca

Introduction

The Barnes maze is a visuo-spatial learning and memory task originally designed for rats and consists of an elevated circular surface with holes around the edge [1]. On the Barnes maze, rats use extra-maze visual cues to locate an escape hole that allows them to escape from open space and bright light into a dark box beneath the maze. The original Barnes maze and test procedure used with rats have been altered in a number of ways for use with mice. As a result, Barnes mazes used with mice differ in size, number of holes, and stimuli used for aversive motivation, while the test procedures differ in the amount of habituation training and the number of acquisition training trials. The present experiment determined if Barnes maze design (maze diameter and presence of a wall and intra-maze visual cues) and test procedure (amount of habituation training) influence learning and memory performance of mice. In particular, we tested whether mice used visuo-spatial cues to learn the location of the escape hole on three different versions of the Barnes maze.

Methods

After receiving either 1 or 4 habituation trials, male (N=30) and female (N=30) C57BL/6J mice were tested on each of three Barnes maze designs; (1) the original design for rats (122 cm diameter), (2) the small maze design (69 cm diameter) and (3) the Pompl maze design (69 cm diameter and a wall with intra-maze cues around the edge) [2]. Mice completed an acquisition phase (15 training days, 2 trials/day) to assess learning, and a probe trial (5 min) to assess memory for the location of the escape hole. The visuo-spatial nature of each maze design was assessed in three ways: (1) a probe trial with a curtain to block extra-maze cues, (2) recording the search strategies used by mice to locate the escape hole and (3) with a reversal test having 5 training days with the escape hole moved to the opposite side of the maze and a probe trial.

Results

The results suggest that the number of habituation trials does not influence learning and memory on the Barnes maze, as performance of mice given one habituation trial did not differ from mice given four habituation trials. Apparatus design did influence maze performance, however, as mice on the large and small mazes spent more time near the escape hole than mice on the Pompl maze during probe trials with visible extra-maze cues. Blocking extra-maze cues during the probe trial with a curtain disrupted performance of mice on the large and small mazes, but not on the Pompl maze. Mice on the large and small mazes also showed greater latency and error reversal effects than mice on the Pompl maze after the escape hole was moved to the opposite side of the maze. Finally, mice on the large maze used the spatial search strategy, which requires the use of extra-maze visual cues to locate the escape hole, whereas mice on the Pompl maze used the serial strategy, which does not require extra-maze cue use. Mice on the small maze used both the spatial and serial search strategies.

Conclusions

These results show that apparatus design influences performance of mice on the Barnes maze. The large maze is the optimal Barnes maze design for testing visuo-spatial learning and memory in mice, as mice reliably use extra-maze visual cues and a spatial search strategy to locate the escape hole. The small maze is a less effective test of visuo-spatial learning and memory than the large maze, as mice on the small maze use extra-maze visual cues, but use both the spatial and serial search strategies to locate the escape hole. The Pompl maze design does not appear to be a visuo-spatial learning and memory test, as mice do not use extra/intra-maze visual cues nor do they use a spatial search strategy to locate the escape hole. Mice on the Pompl maze may not use visual cues because the wall around the edge of the maze promotes thigmotaxis and use of the serial search strategy.

References

1. Barnes, C. A. (1979). Memoy deficits associated with senescence: a neurophysiological and behavioral study in the rat. *Journal of Comparative and Physiological Psychology*, **93**, 74-104.
2. Pompl, P. N., Mullan, M. N., Bjugstad, K. and Arendash, G. W. (1999). Adaptation of the circular platform spatial memory task for mice: use in detecting cognitive impairment in the APP_{sw} transgenic mouse model for Alzheimer's disease. *Journal of Neuroscience Methods*, **87**, 87-95.

Combined effects of illumination, closed wall type and extramaze space size on the anxiety-related behavioral baseline of rats submitted to the elevated plus-maze

N. Violle, F. Balandras, Y. Le Roux, D. Desor, and H. Schroeder

URAFPA, INRA UC340, Nancy-Université, Vandœuvre-lès-Nancy, France, Henri.schroeder@scbiol.uhp-nancy.fr

Introduction

The elevated plus-maze (EPM) is currently used to study the neural basis of anxiety as well as to assess the anxiety-modulating activity of various pharmacological agents. Although the anxiogenic stimulus is supposed to be the open space in this test, the anxiety-related behavioral baseline of rats is sensitive to numerous methodological-linked parameters [1]. In order to improve the quality of behavioral measurements, some modifications have been performed on both original apparatus or procedures initiated by Pellow et al. [2]. Two modifications are often reported: 1) rats are tested under a red light in order to mimic the dark phase of the cycle, and 2) some laboratories used a modified EPM with translucent closed walls to increase the accuracy of the behavioral scoring. At present, the combined effects of these two modifications on spontaneous rat behavior in the EPM have not been sufficiently investigated. Besides, the effect of extramaze space size (i.e. size of the experimental room) on this baseline remains untested. However, an inadequate behavioral baseline may reduce the efficacy of the EPM for measuring anxiety variations in animals. Thus, this study aimed to clarify the respective effects of illumination (ILL), closed-wall type (CWT) and extramaze space size (ESS) on the anxiety-related behavioral baseline of rat in the EPM.

Material and methods

The experiment was carried out on 80 naïve Wistar male rats (Harlan, Gannat, France) maintained in a regulated environment (20±1°C; humidity 50±5%; free access to food and water) under a reversed light/dark cycle (light on at 7 p.m.). Two EPMs were used: a Plexiglas one with translucent walls and a wooden one with opaque walls. They presented identical dimensions and configuration: two opposed open arms (40 x 10 cm) and two opposed closed arms with walls (40 x 10 x 40 cm) linked by a central square (10 x 10 cm) and elevated at 70 cm of height. The EPMs were placed at the centre of a small (2 x 1.7 m) or a spacious experimental room (3.9 x 3.5 m). These two rooms presented identical floors, ceilings, and uniformly white-painted walls. They were empty except for the EPM and the recording facilities. The illumination of the rooms was provided by either 2 white bulbs or 4 red bulbs. The light was homogeneously fitted at 40 lux of intensity by using a luxmeter in such way as the maximal difference observed between the different parts of maze never exceeds 5 lux. Thus, animals were randomly allocated to the 8 experimental groups (10 rat/group) defined by the combination of the 3 tested factors (see table 1). For the testing, each rat was placed at the centre of the maze, head turned toward an open arm, and let freely explore for 5 min. Testing was recorded with a video camera. The number of closed arm entries, the number of open arm entries and the time spent in the open arms were scored from videotapes. The percentage of entries in the open arms and the total number of entries were also calculated. The rat was considered in a part of the maze when is 4 paws were.

Table 1. *Allocation of the 80 rats in the 8 experimental groups defined by the combination of the three tested variables: closed wall type, illumination and extramaze space size.*

Closed wall type	Illumination	Extramaze space size: Small	Spacious
Opaque	Red	n =10	n =10
	White	n =10	n =10
Translucent	Red	n =10	n =10
	White	n =10	n =10

Results

The three-way ANOVA detected significant main effects of ILL, CWT and an interaction between the ESS and the CWT, on behavioral items related to anxiety, i.e., the number open arm entries, the time spent in the open arms (see Figure 1) and the percentage of entries in the open arms. Data analysis indicates that exposure to the white light induced decreases of these 3 behavioral items referred to the red light. Rat behavior presented the same variations when tested in the EPM with opaque closed walls compared to the one with translucent walls. *Post-hocs* also indicates that this closed wall effect was higher in the spacious experimental room than in the small one. Regarding the number of closed arm entries, the statistical analysis showed only a main effect of the ESS which resulted in a significant reduction of this variable in animals tested in the spacious room. Finally, ANOVA indicates main effects of the three factors with interactions between CWT and ESS and between ILL and ESS on the total number of entries.

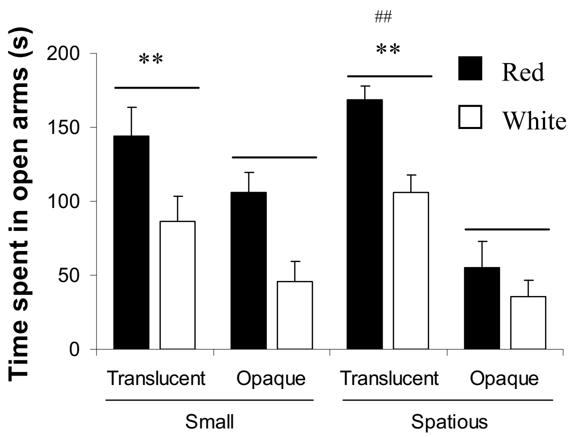


Figure 1. *Effects of illumination (red vs white), closed wall type (translucent vs opaque) and extramaze space size (small vs spacious) on the time spent in the open arms. Data are presented as mean ± S.E.M. Three-way ANOVA detected main effects of illumination and closed wall type (not showed). Tukey's multiple comparisons test was used to study the interaction between closed wall type and extramaze space size: ** p<0.01 different from rats tested in the EPM with translucent walls in the spacious extramaze space. ## p<0.01 different from rats tested in the EPM with opaque walls in the small extramaze space.*

Discussion

Our findings clearly indicates that environmental parameters can strongly modify the animal behavior in the EPM. Concerning ILL, the use of a white light instead of a red one increased the open space avoidance whatever the CWT and ESS, probably due to a less perception of the light intensity when a red one is used [3]. Whatever the type of light used (red or white), the EPM with opaque walls has been perceived as more protective than the translucent one whereas the light intensity was the same everywhere in the maze. Moreover, we founded that the spacious ESS enhanced the effect of the CWT. This result suggests that a spacious ESS may increase the anxiogenic feature of the open arms, but this effect would be more efficient when the closed arms are well discriminated (i.e. opaque walls). Thus, the rat anxiety-related behavioral baseline in the EPM seems to be the result of the perception and the integration of multiple environmental stimuli and to depend on the level of discrimination between the aversive feature of the open space and the relative protection offered by the closed arms. In conclusion, the choice of adequate ILL, CWT and ESS for EPM testing is critical to obtain a behavioral baseline in rats that allowed the detection of an anxiolytic drug. These conditions of testing were applied to the assessment of the activity of the well-known anxiolytic agent, diazepam.

References

1. Rodgers, R.J., Dalvi, A. (1997) Anxiety, defence and the elevated plus-maze. *Neurosciences and Biobehavioral Reviews* **21**, 801-810.
2. Pellow, S., Chopin, P., File, S.E., Briley, M. (1985) Validation of open/closed arm entries in an elevated plus-maze as a measure of anxiety in the rat. *Journal of Neuroscience Methods* **14**, 149-67.
3. Jacobs, G.H., Neltz, J., Deegan, J.N. (1991) Retinal receptors in rodents maximally sensitive to ultraviolet light. *Nature* **353**, 655-656.

Detecting, measuring and correcting errors in automated behavior analysis equipment

Richard E. Brown
Psychology Department, Dalhousie Universit, Halifax, Nova Scotia, Canada B3H 4J1, rebrown@dal.ca

Behavioral phenotyping of transgenic and mutant mice depends on the observation and recording of behavior of specific strains of mice in a particular test apparatus in a particular laboratory [2]. Not all researchers, however, obtain the same results when the same strains of mice are tested in the same apparatus in different laboratories [11]. Chesler et al. [4] calculated that experimenter effects were responsible for a significant proportion of the laboratory effects in behavioral research. These experimenter effects include observer bias, observational errors, and recording errors [7]. In order to reduce experimenter error and conduct high throughput analysis of behavior, many different types of automated behavioral test apparatus have been developed [10].

Automated activity recorders, video-tracking systems, and scoring equipment has been designed for a number of behavioral tests, including the open field, Morris water maze, fear conditioning, and home cage activity. When a researcher uses such automated equipment, how do they know that the data collected are valid and reliable? How can the accuracy of automated equipment be measured?

One method is to compare the scores of automated equipment with those of trained observers. Fitch et al. [5] compared their automated movement detector for fear conditioning with the results of trained observers using three different scoring methods to record freezing behavior. Although all four methods gave comparable results, the automated recording apparatus seemed less sensitive to freezing than event-recorder or time-sampling methods, and event recording allowed the observer to quantify other behaviors. Although Fitch et al. [5] favour of the automated system, the event-recorder and time-sampling methods may provide more information and can be just as accurate. The problem is that they are more time consuming. Crowley et al. [3] found that the scores from automated forced swim and tail suspension tests correlated highly with the scores of three observers but argue that the cost-benefit analysis indicates an advantage of the automated system.

In their analysis of freezing behavior in fear conditioning, Marchand et al. [8] found high correlations (r= 0.995) between observers and between observers and the SUB automated scoring system. However, the two automated scoring methods (SUB vs RAW) produced different results and made different types of errors, classifying some bouts of walking, rearing, sniffing, grooming and moving behavior as “freezing”. Automated equipment is also used to detect more complex behaviors. Graziano et al. [6] calculated that their automated system correctly categorized 97.9% of swim path types in the Morris water maze, as determined by three human observers, with 22 errors in 1049 swim path analyses. Nadler et al. [9] found that two trained observers had a 95% agreement and did not differ in their scores from the automated equipment in scoring sociability (time spent near a target mouse) and social novelty (time spent near a familiar vs. a novel mouse).

In our research, we have compared the results from trained observers and videotape analyses with those of automated equipment and found a number of discrepancies between the data from the automated equipment and the observers. I shall discuss six examples:

Automated open fields give higher activity scores than observer-based open fields. In testing Coloboma mice we found that the automated open field gave very high “horizontal activity” scores because circling was scored as horizontal activity and not as stereotyped behavior.

Automated open fields track only part of the mouse, not the whole mouse. When scoring activity, we score a movement only when all four feet of the mouse cross a line, but our automated system scored only the front half of the mouse, thus transition scores were inflated.

The automated Barnes maze made errors in scoring“head pokes” into holes. Our tracking system defined a zone around the hole as an “error zone” and scored an error when a mouse entered this zone, even if it did not head-poke into the hole. Thus, the automated system scored far more errors than the human observers.

Automated tracking system errors. We found tracking system errors on the elevated plus maze, open-field and Morris water maze as the tracking system recorded behavior “outside the apparatus” and showed the mouse travelling through the walls of the apparatus or “flying” from one arm of the elevated plus maze to another. We corrected these errors manually. They appear to be due to miss-alignment of the equipment or lights.

The five-choice serial reaction time box was not designed so that mice could meet the criteria for learning. We are currently redesigning the apparatus to eliminate this problem and this will allow us to complete the experiment in less than half the time required in published papers.

The automated recording of freezing behavior in cued and context conditioning resulted in errors in setting the baseline and in recording freezing as discussed by Marchand et al. [8].

The discrepancies between the automated equipment and the observers were due to (1) different operational definitions of the behaviors, (2) equipment hardware and software problems, and (3) improper adjustment of the equipment. We have corrected the hardware, software and equipment adjustment problems but the problem of operational definitions remains an issue.

One of the problems with behavioral phenotyping is that researchers often receive no training in behavior analysis [1]. With the pressure to test more and more animals faster and faster, automated equipment will be used more often, but, while this equipment increases the speed of testing and compares favourably with trained observers on latency measures, the error rate of automated equipment is often unknown. Because test apparatus is not standardized, new apparatus is often not tested parametrically for reliability and validity, and if experimenters are not trained observers of behavior, the equipment errors go undetected. In experiments using automated equipment, we recommend that exact details be given in the methods section of papers, behaviors be videotaped and that the results from automated equipment be verified by trained observers and that error rates for such equipment be calculated.

References

1. Blizard, D.A., Takahashi, A., Galsworthy, M.J., Martin, B., Koide, T. (2007). Test standardization in behavioural neuroscience: a response to Stanford. *Journal of Psychopharmacology* **21**, 136-139.
2. Brown, R.E. (2007). Behavioural phenotyping of transgenic mice. Canadian *Journal of Experimental Psychology* **61**, 328-344.
3. Crowley, J.J., Jones, M.D., O'Leary, O.F., Lucki, I. (2004). Automated tests for measuring the effects of antidepressants in mice. *Pharmacology, Biochemistry and Behavior* **78**, 269-274.
4. Chesler, E. J., Wilson, S. G., Lariviere, W. R., Rodriguez-Zas, S. L., Mogil, J. S. (2002). Identification and ranking of genetic and laboratory environment factors influencing a behavioral trait, thermal nociception, via computational analysis of a large data archive. *Neuroscience and Biobehavioral Reviews* **26**, 907-923.
5. Fitch, T., Adams, B., Chaney, S., Gerlai, R. (2002). Force transducer-based movement detection in fear conditioning in mice: a comparative analysis. *Hippocampus* **12**, 4 -17.
6. Graziano, A., Petrosini, L., Bartoletti, A. (2003). Automatic recognition of explorative strategies in the Morris water maze. *Journal of Neuroscience Methods* **130**, 33-44.
7. Lehner, P.N. (1996). *Handbook of Ethological Methods*. 2nd ed. Cambridge: Cambridge University Press.
8. Marchand, A.R., Luck, D., DiScala, G. (2003). Evaluation of an improved automated analysis of freezing behaviour in rats and its use in trace fear conditioning. *Journal of Neuroscience Methods* **126**, 145-153.
9. Nadler, J.J., Moy, S.S., Dold, G., Trang, D., et al. (2004). Automated apparatus for quantitation of social approach behaviors in mice. *Genes, Brain and Behavior* **3**, 303-314.
10. Tecott, L. H., Nestler, E. J. (2004). Neurobehavioral assessment in the information age. *Nature Neuroscience* **7**, 462-466.
11. Wahlsten, D., Metten, P., Phillips, T. J., et al.. (2003). Different data from different labs: Lessons from studies of gene-environment interaction. *Journal of Neurbioogy* **54**, 283-311.

Telemetric Longitudinal Measurement of Young Driver Behaviour

Nicola J. Starkey[1] and Robert B. Isler[2]
Traffic and Road Safety Research Group, Department of Psychology, University of Waikato, Hamilton, New Zealand
[1]nstarkey@waikato.ac.nz; [2]psyc2255@waiakto.ac.nz

New Zealand is one of the few countries that currently allows teenagers to become solo drivers at 15 ½ years of age. During their first six months of solo driving, these teenagers are 19 times more likely to crash compared to the period of supervised driving. The behaviour of these adolescent drivers represents the single largest cause of fatalities in that age group and is widely acknowledged as one of the most serious social issues facing New Zealand. This state of affairs, often referred to as 'the young driver problem,' with an over representation of young drivers in motor vehicle crashes, is not unique to New Zealand. In fact, a similar situation can be found inmost of the world's developed nations.

Contributing to this problem, traditional methods of driver education and driver training have not delivered demonstrable safety benefits. In fact, the majority of driver training evaluation studies in the last thirty years have concluded that driver education and training contributes little to reduce crash risk / involvement for road users (pre-licence, defensive, advanced, or driver improvement). In fact, some types of car handling training has been cited as leading to increased risk taking due to learners' inflated self confidence and self rated driving skills [1]. However, accurately evaluating the effects of driver training interventions is a difficult task. A particular driver training intervention can only be considered to be effective if it can show a significant reduction in the number crashes for the driver, or a significant change in driver behaviour that clearly implies safer driving. Getting accurate and comprehensive crash records is difficult and to measure post training behavioural driving changes based on self-reports (e.g., log books) may not be accurate enough to be statistically meaningful. In fact Crick and McKenna (1991) suggest that the lack of evidence for the benefits of road safety education / training may be due to a lack of methodological soundness in previous evaluations [2].

Event data recorders have been used for years in aircrafts and more recently in cars to monitor speed, seat belt use and air bag release after a crash. In addition they are used in vehicle fleets to track location and provide information on risky driving. More recently, this technology has become widely available for purchase and can be installed in cars once they have left the factory, providing information on speed, distance travelled, location and large g force changes. Such devices have the potential to be used in a variety of ways to improve driver safety and education. Firstly, the long term effects of driver education or training could be accurately evaluated. Secondly, the behaviour of young drivers could be monitored and unsafe driving practices could be highlighted. Indeed, with the addition of a video camera this provides opportunities for driver education [3]. Finally, these devices could be used to improve fuel consumption and decrease risky driving [4].

In terms of monitoring unsafe driving practices among teen drivers, several 'black box' devices are available for purchase over the internet (e.g. RS-1000 Teen Black Box – see Figure 1, Alltrackusa). Data stored on the devices can be downloaded to a PC and reviewed by parent / driver of the vehicle. In addition, audible beeps alert the driver to unsafe driving as it happens. Thus, information from these devices allows parents to monitor how their children are driving but gives little information regarding the situational factors which led to unsafe driving.

To take advantage of the educational opportunities this type of monitoring provides, McGhee et al, 2007 added an event triggered video device to the black box. The 'DriveCam' captures video continuously, providing a forward and interior view of the car. Data is only stored (20 seconds) when a preset accelerometer threshold is exceeded. Events can then be downloaded from the device and are coded in the laboratory. Each week, the parent and teen receive a CD to review (showing the recorded events) with a report card and suggestions on how to improve their driving [3].

***Figure 1.** The RS1000 Teen Black Box*

Our research is focused on improving driver education and as part of this we carried out a pilot study using a telemetric data tracker to determine how well this technology measured real driving behaviour. After a driver training camp we installed telemetric data trackers in the vehicles of eight participants to pilot how well this technology measured post-training real driving behaviour over a period of 32 weeks. The tracking system consisted of a small credit card sized global positioning module (SmarTrak Lite GPRS / GPS) fitted with an accelerometer. The software for the tracking and reporting interface via the internet was developed by SmarTrak Ltd (www.smartrak.co.nz). It allowed us to monitor, in real time, the driving performance (updated every 2 seconds) of the eight participants on the computer screen (see Figure 2). The built-in accelerometer also provided g-force data from the vehicles. Daily, weekly and monthly reports of the driving measures for each participant could be produced and downloaded as a Microsoft Office EXCEL spreadsheets. For each participant we obtained d*istance driven per trip;* n*umber of trips; mean speed per trip;* m*aximum speed; speeding violation and large G-force*

We received valid telemetric driving behaviour data from six of the eight participants. Two of the participants crashed during the study and the GPS system allowed us to examine their driving behaviour just before (and, in one case, during and after) the crash. In addition, one participant had their car stolen. The data we received provided us with an interesting insight into the driving patterns of these teenagers and overall, the telemetric data tracking system used in this study seems to be a promising research tool for evaluating driver behaviour.

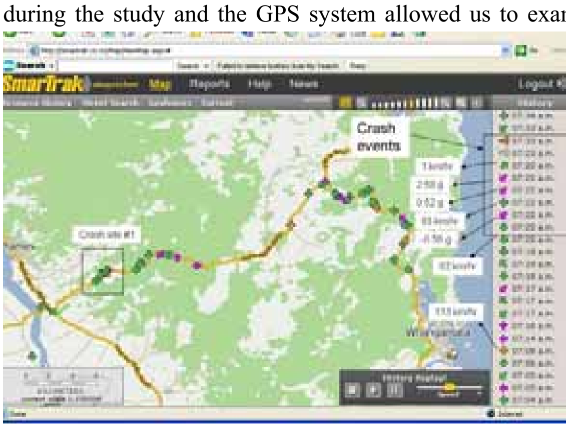


***Figure 2.** The map function of the on-line monitoring system*

By using the map based tracking function all the recorded driver behaviour events, including crashes could be mapped, replayed and analysed in detail on the internet. It also allowed us to create daily, weekly and monthly reports of important risk-taking behaviour variables (such as speeding, average speed, large g-forces) and could also provide information on risk exposure (driving distance). In order to improve the system, an event triggered video recording system could help verify each large g-force that was created by the monitored vehicles. It would also be beneficial to record lower speeding events such as driving 60 km/h on a road with a 50 km/h speed limit, but this depends on GPS based speed limit data for all roadways being available. The difficulties associated with this technology include the huge amount of data which is obtained and also installing these devices in private cars raises ethical issue relating to privacy. Overall this technology appears to provide a sensitive and reliable means of evaluating driver behaviour for a range of purposes.

References

1. Gregersen, N.P. (1996). Young drivers' overestimation of their own skill – An experiment on the relation between training strategy and skill. *Accident Analysis & Prevention* **28** (2), 243-250.
2. Crick, J.,& McKenna, F.P. (1991). Hazard perception: Can it be trained? In G.B. Grayson (ed.) Behavioral Research in Road Safety II. Proceedings of a seminar at Manchester University 17-18 September 1991. Crowthorne, Berkshire: Transport Research Laboratory. pp.100-107.
3. McGehee, D.V., Raby, M., Carney, C., Lee, J.D., and Reyes, M.L. (2007). Extending parental mentoring using an event–triggered video intervention in rural teen drivers. *Journal of Safety Research* **38**, 215-227.
4. Greenroad (2007) Whitepaper downloaded from http://www.greenroadtech.com/documents/fuel-consumption.pdf.

Analysing the influence of driving experience on the glance behaviour at selected urban crossroad types

M. Williams
Institute of Ergonomics, Darmstadt University of Technology, Germany, williams@iad.tu-darmstadt.de

Usually when analysing the influence of the driving experience on his/her behaviour parameters like the total driving experience, i.e. since getting the driving license or the average driving experience i.e. total driving experience divided by the number of years (see also [1]). In contrast less attention is paid to other components like for example the driving experience accumulated in the last year or number of kilometres in urban traffic. In this study it was analysed to what extent these components of the driving experience are linked with the glance behaviour during selected crossing situations in urban traffic.

Experimental setting

In this section the selected crossroads, test subjects, and test vehicle will be described.

Selected crossroads

The main selection criterion consisted in the demands posed to the driver. The demands associated to a driving situation concern both information processing and vehicle control [2]. The selected crossroads had to pose different demands to the driver, i.e. in one case low demands and in the other one high demands. Two urban crossroads have been selected for this experiment; at the first one the traffic was controlled by traffic light (see figure 1) whilst the second crossroad was sign-posted (see figure 2).

Figure 1. *Crossroad with traffic light*

Figure 2. *Sign-posted crossroad*

Whilst approaching the first crossroad the traffic light switches to red so the driver has to stop and wait for green before crossing; the demands posed to the driver are low. The approaching phase started when the traffic light become visible for the driver. At the second crossroad the driver has to give way and to wait for an adequate gap in order to cross; the demands posed are high. The approaching phase started here 3 seconds before reaching the waiting line. Both approaching and waiting have been considered in this study.

Test subjects

Thirty subjects (18 males, 12 females) participated in the experiment. The youngest test subject was 20 and the oldest 67 (Mean = 45.8, std. dev. = 17.8 yrs). Following components of the driving experience have been considered: total driving experience, driving experience accumulated in the previous year, driving experience in urban traffic (% from the total driving experience). The minimum total driving experience was 12,000 km and the maximum 1,200,000 km (Mean = 405,833.3 km, Std. Dev. = 335,697.5 km). In the previous year there were driven between 6,000 and 50,000 km (Mean = 17,400 km, Std. Dev. = 10,040.8 km). Besides, between 10 and 80% of the total driving experience was gathered in urban traffic (Mean = 36.2%, Std. Dev. = 17.6%).

In order to be able to analyse the differences between the test subjects regarding the glance behaviour different groups have been built (see Table 1).

Table 1. Driver groups

Driver characteristic	Groups		
Total driving experience (km)	< 100,000 N = 7	100,001 – 500,000 N = 17	> 500,001 N = 6
Driving experience in the previous year (km)	5,001- 10,000 N = 9	10,001 – 20,000 N = 12	> 20,001 N = 9

The experience in the urban traffic was considered as a continuum; there were no groups built.

Proceedings of Measuring Behavior 2008 (Maastricht, The Netherlands, August 26-29, 2008)
Eds. A.J. Spink, M.R. Ballintijn, N.D. Bogers, F. Grieco, L.W.S. Loijens, L.P.J.J. Noldus, G. Smit, and P.H. Zimmerman

Test vehicles

Two test vehicles have been used, both equipped with cameras for recording the glance behaviour of the driver and with sensors for collecting dynamical data, i.e. speed and acceleration.

Methodology for data analysis

In a first step the relevant phases (i.e. approaching and waiting) for both situations have been identified and marked. The Observer© 3.0 has been used for the analysis of the video tapes. Consequently, a structure has been programmed; it contained following behavioural classes: situation, street type, glance direction (i.e. forward, to the left, to the right, to the traffic light, other directions), and presence of other road users ahead. For the sign-posted crossroad an additional behavioural class has been included: road user entering the crossroad from the left / from the right.

But The Observer delivers for each direction only the duration and the frequency; which is not enough in order to get a differentiated picture of the glance behaviour. Because the durations of the approaching and of waiting are different from one test subject to another it was necessary to perform normalization, i.e. both frequency and duration have been divided by the particular phase duration. In this way a base for comparison has been set.

Besides, in order to better describe the visual attention paid by the test subjects while waiting to the road users entering the sign-posted crossroad, both frequency and duration have been divided by the number of road users entering the crossroad. The different analysis levels are presented in Table 2.

Results

The results showed that the scheme is adequate in order to study the glance behaviour at crossroads posing different demands to the driver; it delivers a differentiated picture of the glance behaviour. Two components of driving experience, namely total driving experience and the experience gathered in the previous year led to significant differences between driver groups with regard to the glance behaviour. Although urban crossroads have been selected, the experience in urban traffic plays only sporadically a role.

Table 2. *Analysis scheme for the glance behaviour*

Glance Direction	Frequency	Density = Frequency / Phase duration
		Frequency / road user = Density / Number of road users*
	Duration	Percentage from phase duration = Duration / Phase duration
		Average duration = Duration / Frequency
		Duration / road user = Duration / Number of road users*
Level I	**Level II**	**Level III**

* *Only for the sign-posted crossroad.*

References

1. Breuer, J. (1996). Ergonomische Beurteilung und Gestaltung der Sicherheit des Arbeitssystems Kraftfahrzeugführen. *Fortschritt-Berichte VDI*. Reihe 12. Nr. 271. Düsseldorf : VDI Verlag.
2. Fastenmeier, W. (1995): Die Verkehrssituation als Analyseeinheit im Verkehrssystem. In: *Fastenmeier, W. (Hg.): Autofahrer und Verkehrssituation. Neue Wege zur Bewertung von Sicherheit und Zuverlässigkeit moderner Straßenverkehrssysteme. Reihe Mensch – Fahrzeug – Umwelt.* Köln : Verlag TÜV Rheinland. 27–78.

The *e-motion* system: an integrated device to evaluate drivers' functional state in the field

A. Clarion[1,2], C. Ramon-Zarate[3], C. Petit-Boulanger[1], A. Dittmar[3] J.P. Bourgeay[4], and C. Collet[2]

[1]*Ergonomics and Human Factor Research Department, Renault, Guyancourt, France, antoine.clarion@renault.com*

[2] *Mental Processes and Performance Laboratory, Claude Bernard University Lyon 1. Villeurbanne, France, christian.collet@univ-lyon1.fr*

[3] *Microsensors and Biomedical Micro systems, INSA de Lyon, Villeurbanne, France. carolina.ramon-zarate@insa-lyon.fr*

[4]*By-Consulting Society, Saint Genis Laval, France, by-consulting2@wanadoo.fr*

Research has now provided ample evidence that drivers' functional state should be assessed through multiple recordings from several fields. Drivers' functional state may be studied indirectly through data from the vehicle. Physiological indicators are believed to bring reliable information when available directly and often give anticipated cues about the forthcoming behaviour. A set of bio signals such as electrodermal activity, heart rate or respiratory frequency represent useful and objective data which may be recorded in real time. Indeed, variables of the autonomic nervous system are known to mobilize energy resources of the organism, in response to internal and external milieu demands. The main role of the sympathetic branch is to face up emergency situations. As electrodermal activity (EDA) is under the control of the sympathetic nervous system only, it is closely correlated to arousal and its variation. Recordings from the vehicle and the driver himself through specific systems integrating data on the same time scale should give valuable information for a close estimation of drivers' functional state, its evolution across time and thus to a better understanding of human being functioning.

Many separated systems do exist, each with a specific aim at studying a particular aspect of drivers' behaviour. The main objective of our work is to propose an integrated system, made of several sub-units, each capable of giving valuable solution to a scientific question related to drivers' behaviour.

Specific needs

There is a need to coordinate data from the car to information from the driver himself (physiological and behavioural parameters) and to external data from the environment. Dynamic vehicle parameters provide objective data (speed or command uses). More, some of these parameters give direct information about driver's motor activity (e.g. steering wheel angle or pedal depressions). Physiological analogical signals are recorded from micro sensors placed on the non-dominant hand (the palm is strongly innervated by the sympathetic skin endings, controlling sweat glands). Signals could be processed before being digitized via an acquisition card and analyzed with specific software. Video is aimed at providing data about the driving context and the in-vehicle scene. Data could be compressed easily and recorded on hard disk and bring information from four different scenes (front and rear driving space, drivers' feet position on the pedals and large plan on the driver's face).

Thus, there is a specific requirement to provide synchronization of psycho physiological, video/contextual and vehicle parameters recordings. Such a device must also be car-embeddable. The last stage is to benefit from data files easy to read and analyze with a dedicated and ergonomic tools.

Vehicle parameters: CAN

A CAN bus is available on each car and the CANalyser system (Vector GmbH) is aimed at recording data from the main car systems. These are stored as a CANalyser file, available for further exportation prior to analysis.

Physiological signals recording device

The recording device has the potentiality to record several different physiological signals: skin resistance, skin potential, raw ECG, or nasal air flow temperature. Any other parameter may be selected if needed (skin temperature, skin blood flow). This acquisition device is made to convert analogical voltage measures into relevant digital physical quantities, sampled at 10 Hz. A dedicated software interface called *e-motion* is dedicated to pilot the device. Physiological data are then stored in dedicated files: a new file is created and saved every 5 minutes to prevent accidental data loss.

Video

A quad unit concatenates video flux from four analogical cameras and a frame grabber (Dazzle DVC 80) converts video signal from analogical to digital. The video acquisition software, called *V-motion* compresses video data at a DivX format and stores data in *avi* file-format. The *V-motion* software makes the slaves storing data in the same file directory. The experimenter may have a look on the recordings at any time, date each event by using an event-marker and add any commentary related to the observation of the road scene. Further analysis may then associate cinematic and physiological data with each event and thus bring objective information related to drivers' behaviour.

Contextual data

External events may easily be tagged during data acquisition. The more probable events may be previously predicted and thus pre-written in a specific configuration file (using the *PiloteErgo* software). Unexpected events may nevertheless be added easily during the experiment.

Data synchronization

The *V-motion* software integrates a strip with the time elapsed since start. A reliable time-scale is easily obtained from the core's computer. This time is given as the reference time for all other devices. As shown by Figure 1, the *V-motion* device is the master software that triggers and stops all other acquisitions, including the control of the *e-motion* device driver program, the *CANalyser* control via the *PiloteCAN* software and the setting of *PiloteErgo*'s time for contextual events dating.

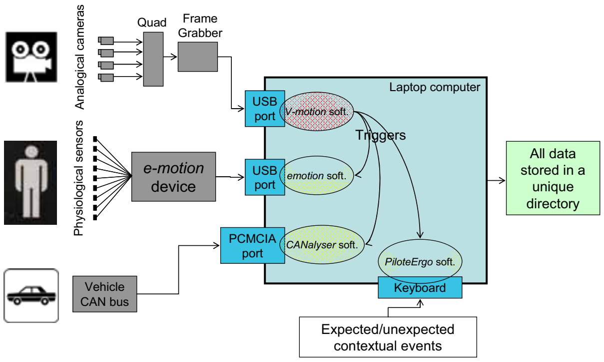


Figure.1. *Schematic representation of the data acquisition device.*

Data analysis

Data analysis is based upon the *LVA* and the *CaroLab* sofware. The *LVA* manages the video movie while the *CaroLab* (*MatLab* - The Mathworks, Inc.) provides the interface displaying both physiological and CAN data with data processing tools. Both display contextual information, thus providing the experimenter meaningful information (particular attention is given to physiological variation related to contextual information. The *CaroLab* software is designed to select any contextual event, to display the events of the same category with the same coloured vertical line and to change the framing as needed.

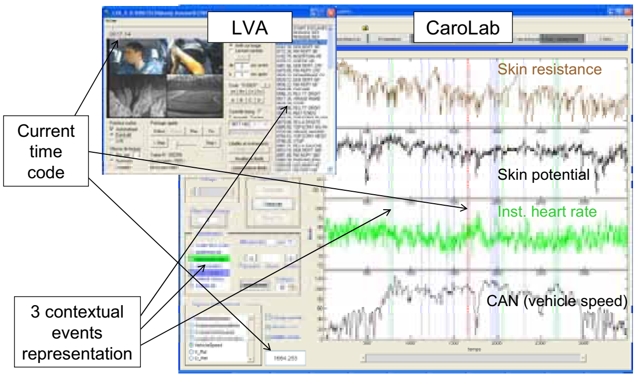


Figure.2. *Schematic representation of the data screens available on the same time-scale.*

Signal Processing

The CaroLab software provides several semi-automatic data processing tools. A butterworth smoothing filter can be applied on raw data before analysis. The software is designed to compute several indicators around each contextual event and to store data in a specific worksheet. The program sets 2 cursors for automatic response detection. The user can move them if the detection does not fit well. Physiological response amplitude, duration, slope, latency, onset value and mean value is thus easily obtainable and associated with behaviour. Future tools should further be developed e.g. heart rate variability and Poincaré plots.

Conclusion

This integrated system provides an effective data acquisition solution, available from field tests in a car-embedded context. During the experiment, any useful contextual information may be tagged in real time to be associated with behaviour and physiological activity attached to it. Thus drivers' behaviour may be assessed through different indicators. This multivariate analysis brings redundant information and is thus aimed at improving reliability.

Leadership behavior of Dutch primary school leaders: How does their behavior contribute to the effectiveness of their schools?

J.H.E. Nijhuis[1] and C.P.M. Wilderom[2]

[1]*University of Twente (Information Systems & Change Management), Faculty of Management & Governance, and Financial Controller, Stichting Consent, j.nijhuis@consent-enschede.nl*

[2]*University of Twente (Information Systems & Change Management), Faculty of Management & Governance*

Introduction

School leaders are held accountable for their student achievements even though most studies find that school leadership does not have a direct effect. Our study tested a number of hypotheses that relate primary school leadership to student achievement in indirect ways. In addition to employing a survey, we observed the real-life behavior of highly effective school leaders and less effective school leaders; we filmed the behavior of 20 primary school directors/leaders, each in one pre-scheduled, regular meeting with their teachers. A new coding scheme based on extant (educational) leadership theory plus the behavioral software "The Observer" enabled us to analyze the films, sentence for sentence. Our results further detail as well as go beyond the transformational, transactional and educational leadership paradigm.

The eleven exclusive behaviors we have defined and coded/analysed were mostly based upon the Multifactor Leadership Questionnaire from Bass & Avolio [1], House [2,3] and Pearce [4]. These observed behaviors are further improved by Wilderom & Van der Weide [5,6]:

1. showing disinterest
2. defending one's own position
3. providing negative feedback
4. directing
5. verifying
6. structuring the conversation
7. informing
8. visioning
9. intellectual stimulation
10. individual consideration
11. active listening

Although one of the main reasons for conducting this video-observation study was the indirectness of questionnaire studies, behavioral observation studies might benefit from combining video-taping methods with the more traditional leadership research methods like questionnaires. Hence, apart from our behavioural videotaping, we surveyed perceived educational leadership in terms of the following three dimensions, based upon the studies of Hendrikse, Doolaard, Lam & Bosker [7]:

1. scheduled meetings (we used 11 items)
2. professionalizing (we used 14 items)
3. educational leadership (we used 14 items)

The research into transformational leadership in educational settings was initiated by Leithwood [8,9,10,11,12] and his colleagues from the Ontario Institute for Studies in Education (Toronto, Canada) in the late eighties and early nineties. Leithwood's survey research concerns not only the nature of transformational school leadership but also the internal process of transformational school leaders and the effect of such leadership on the school. With regard tot internal processes of transformational school leaders, Leithwood & Stager [13] report high levels of problem-solving expertise. The results of Leithwood's studies into the nature of school leadership (based on the aformentional work of Burns and Bass) have revealed specific dimensions of transformational school leadership and behaviors associated with each of these dimensions. In addition we employed these three dimensions of transformational leadership, as they were the most relevant in our filmed context. The survey items we used are based upon the studies of Sleegers & Geijssel [14].

1. Charisma/inspiration/vision which means inspiring teachers to be engaged in their work by developing, identifying, and articulating a particular vision;
2. Individual consideration which means concern and respect for the personal feelings and needs of teachers; and
3. Intellectual stimulation which means challenging teachers to professionalize themselves in such a manner that the organization is learning as a whole

Data collection issues

A data-collection issue in the video-study pertains to the perspective of the camera. In principle there are two possibilities: the etic and the emic view [15]. The emic view refers to the situation where the researcher is participating in the social context to get a view from the inside-out. The converse, etic view – as used during our study – enables the observation of school leader behavior from the outside. This technique was useful so as to obtain a more objective view of the social reality/ situation. For an etic view study it is important that the video camera is located at a fixed place and forms a part of the background [16,17], thus reducing the level of obtrusiveness [18]. An additional advantage is that all videos are recorded from the same position. Hence, the person operating the camera has less freedom of movement thus keeping his or her intrusion as limited as possible.

Another key issue is that of reactivity. Once people know they are being observed, or as soon they come into contact with one another, there will be a certain degree of so-called 'reactivity.' People behave differently in a group than when they are alone, or as Goffman [19], puts it: "Thus, when the individual presents himself before others, his performance will tend to incorporate and exemplify the officially accredited values of the society, more so, in fact, than does his behavior as a whole." As such, the level of reactivity is affected by the social value of the behavior. Behavior that is negatively valued (e.g., verbal abuse) will occur less frequently during an observation than socially desirable behavior. A research subject can more or less decide to perform specific behavior during the observational period [20,21]. Field experiments still have to consider reactivity, yet, in experiments, subjects are not studied in their 'natural' surroundings. In addition, in a laboratory experiment subjects are more reflective toward their behaviors than in their own habitat.

For a field study like ours there are three most common strategies to minimize any possible reactivity. The first one is unobtrusiveness. Yet for practical and ethical reasons, installing hidden video cameras in multiple organizations for observing multiple managers is a bridge too far. Another strategy is manipulation: by withholding the real purpose of the study from the research subject and offering alternative reasons. However, from an ethical point-of-view it is of utmost importance to correctly inform and treat the research subjects [22,23]. Hence, we promised the school leaders that the material would only be used for scientific purposes. For these reasons we chose to follow the third strategy, which is that of acclimatization. Enough time in the field can make the researcher less visible – and hence non-reactive – in the setting so that the normal flow of activities can resume [24]. Therefore, data has only been collected after the research subject has accustomed to the presence of the video camera. First, we explained to the school leaders and their teachers what we were going to do. Then, right before starting the recording, we met for half an hour with the research participants to get acquainted and to answer any questions. This was all useful so as to help the participants to become familiar with the video camera setting [25].

Furthermore, the role of the researcher needs to be articulated [26]. In our study this was crystallized in the role of the camera operator. Previous studies showed that the role of the researcher was seen as vague and mysterious [26]. We endeavored to ensure that the participants would not attribute a negative role to the camera operator. Also, by being present during the entire meeting, the researcher could gather an in-depth feeling of the meeting and its context.

We chose to use one camera; using multiple cameras would have made the field research far more complex [27]. The camera was always put in the same clearly visible location, to add to a trusting relationship between the researchers and his subjects. According to Kent and Foster [18], behavioral observation procedures, making use of videotaping, usually seem to be largely unobtrusive and unaffected by the expectations of the examiner. The use of the camera is perhaps even less obtrusive than writing down field notes as it causes less anxiety and curiosity among the subjects. A camera does not give feedback to specific behaviors (reciprocity); rather it continuously, carefully, and reliably observes behavior. Furthermore, the camera was solely directed at the school leaders, thus further lowering the reactivity of their "followers." Hence, the role of the researcher in all of the video-recording in this study and the position of the camera were carefully chosen as to have as little reactivity as possible. As school directors and their teachers refrained from asking unnecessary questions related to the filming, our method seems quite unobtrusive.

Data analyses of the video-study

After converting the videotapes into so-called MPEG data files, we were able to codify the behaviors of the school leaders. We used the behavioral-software program 'The Observer,' which saved a considerable amount of time otherwise required for transcription [28,26]. 'The Observer' also allowed us to watch and code the video images at the same time with help of a behavioral coding scheme. These codes were registered and resulted in a separate data file for every one-on-one meeting and department meeting. In an effort to make the coding work easier, we listed the coding scheme of eighteen behaviors in a friendliness order, rating managers' behaviors from very empathic to very self-interested.

Two observers worked on coding the data collected in our sample. These coders were thoroughly trained in the behavioral coding scheme [29]. The thoroughness of the coding is made apparent by the amount of time required: each minute of video required an average of three minutes to code. Two observers separately coded the same video images. The two observers then reviewed the coding differences together [18,30]. We used inter-rater reliability as an indicator of need for standardization of the coding scheme [31,32]. The measure we used to establish inter-rater reliability is the percentage of agreement among the raters (i.e., 'the percentage of raters who agree to a specific code for a given section of the recording').

In our presentation we will link the obtained behavioral results of the study to the survey results: the survey was administered among the attendees of each filmed meeting: right after each filming session.

References

1. Bass, B.M., Avolio B.J. (2000). MLQ Multifactor Leadership Questionnaires Technical Report (2nd ed. Redwood City: Mind Garden.
2. House, R.J., Hanges, P.J., Ruiz-Quintanilla, S.A., Dorfman, P.W., Javidan, M. Dickson, M., et al. (1999). Cultural influence on leadership and organizations: Project GLOBE. *Advances in Global Leadership* **1**, 171-233.
3. House, R.J., Hanges, P.J., Ruiz-Quintanilla, S.A., Dorfman, P.W., Javidan, M. Dickson, M., et al. (2006). Questionnaire GLOBE Project, form Beta. Retrieved May 14, 2007, from http://www.thunderbird.edu/wwwfiles/ms/globe/pdf/GLOBE_Phase_2_Beta_Questionnaire.pdf.
4. Pearce, C.L., Sims, H.P. Jr., Cox, J.F., Ball, G., Schnell, E., Smith, K.A., & Trevino, L. (2003), Transactors, transformers and beyond. A multi-method development of a theoretical typology of leadership. *Journal of Management Development* **22(4)**, 273-307.
5. Van der Weide, J.G. & Wilderom, C.P.M. (2004). Deromancing leadership: What are the behaviors of highly effective middle managers? I*nternational Journal of Management Practice* **1(1)**, 3-20.
6. Van der Weide, J.G. (2007). Een explorative video-observatiestudie naar het gedrag van effectieve middenmanagers. Unpublished doctoral dissertation, University of Tilburg,Tilburg. (Ridderprint, Ridderkerk).
7. Hendrikse, M.A., S. Doolaard, J.F. Lam en R.J. Bosker (2000). Het instrumentarium "school- en klaskenmerken", ontwikkeling en normering (ZEBO-SK). Enschede: Universiteit Twente, Onderzoek Centrum Toegepaste Onderwijskunde.
8. Leithwood, K.A., en Montgomery, D.J. (1986). Improving principal effectiveness: The principal profile. Toronto: OISE.
9. Leithwood, K.A., P.T. Begley en J.B. Cousins (1990). The Nature, Causes and Consequences of Principals' Practices: An Agenda for Future Research. *Journal of Educational Administration* **28(4)**, pp.5-31
10. Leithwood, K.A., Jantzi, D., en Steinbach, R. (1998, April). So school councils matter? Paper presented at the annual meeting of the American Educational Research Association, San Diego, C.A.
11. Leithwood, K.A., Jantzi, D., en Steinbach, R. (1999). Changing leadership for changing times. Philadelphia: Open University Press.
12. Leithwood, K.A., Jantzi, D., (2006), Transformational School Leadership for Large-Scale Reform: Effects on students, teachers, and their classroom practices. *School Effectiveness an School Improvemen*t, **17(2)**, 201-227
13. Leithwood, K. en Stager, M. (1989). Expertise in principals' problem solving. *Educational Administrative Quartely,* **25 (2)**, 126-161.

14. Geijssel F. (2001). Schools and innovations: conditionas fostering the implementation of educational innovations.Nijmegen: University Press.
15. Ratcliff, D. (2003). Video methods in qualitative research. In P.M. Camic, J. E. Rhodes, & L. Yardley (Eds.), Qualitative Research in Psychology. Washington DC: American Psychological Association.
16. Erickson, F. (1992). The interface between ethnography and microanalysis. In M.D. LeCompte, W.L. Millroy, & J. Preissle (Eds.), The Handbook of Qualitative Research in Education. San Diego: Academic Press.
17. Mead, M. (1995). Visual Anthropology in a discipline of words. In P. Hockings (Ed.), Principles of Visual Anthropology. New York: Mouton de Gruyter.
18. Kent, R.N., & Foster, S. L. (1977). Direct observational procedures: Methodological issues in naturalistic settings. In A.R. Ciminero, K.S. Calhoum, & H.E. Adams (Eds.), *Handbook of Behavioral Assessment* (pp. 279-328). New York: John Wiley.
19. Goffman, E. (1959). The presentation of self in every day life. Garden City, NY: Doubleday.
20. Weick, K.E. (1985). Systematic observational methods. In.G. Lindzey & E. Aronson (Eds.), T*he Handbook of Social Psychology (3rd ed.).* New York: Random House.
21. Roth, J. (1989). The Disturbed Subject: Epistemological and Ethical Implications of Reactivity in Videotape Research. New York: Lang.
22. Rosenstein, B. (2002). Video use in social science research and program evaluation. International *Journal of Qualitative Methods*, **1(3)**, 1-38.
23. Schaeffer, J.H. (1995). Videotape: new techniques of observation and analysis in anthropology. In P. Hockings (Ed.), *Principles of Visual Antropology* (2nd ed.). New York: Mouton de Gruyter.
24. Goetz, J.P., & LeCompte, M.D. (1982). Ethnography and Qualitative Design in Educational Research. Orlando: Academic Press.
25. Bourchard, T.J. (1970). Field research methods: Interviewing, questionnaires, participant observation, systematic observation, unobtrusive measures. In J.P. Campbell, M.D. Dunnette, E.E. Lawler, & K.E. Weick (Eds.), Managerial Behavior, Performance and Effectiveness. New York: McGraw-Hill.
26. Ratcliff, D. (1996). Video and Audio Media in Qualitative Research. Retrieved June 6, 2007, from http://don.ratcliff.net/video
27. Jackson, B. (1987). *Fieldwork.* Urbana: University of Illinois Press.
28. Lancy, D.F. (1993).*Qualitative Research in Education: An Introduction to the Major Traditions.* New York: Longman.
29. Reid, J.B. (1982). Observing training in naturalistic research. In P.M. Camic, J.E. Rhodes, & L. Yardley (Eds.), *Qualitative Research in Psychology.* Washington DC: American Psychological Association.
30. Beck, D., & Fisch, R. (200). Argumentation and emotional process in group decision-making: Illustration of multi-level interaction process analysis approach. *Group Processes & Intergroup Relations,* **3(2)**, 183-201.
31. Goldfried, M.R., & Linehan, M. (1977). Basic issues in behavioral assessment. In A.R. Ciminero, K.S. Calhoun, & H.E. Adams, *Handbook of Behavioral Assessmet.* New York: John Wiley & Sons.
32. Van de Sande, J.P. (1986). *Gedragsobservatie. Een inleiding tot systematisch observeren.* Groningen: Wolters-Noordhoff.

Towards an Annotation Standard for Eye Tracking Data

Michael Carl[1], Arnt Lykke Jakobse[1], and Oleg Spakov[2]

[1]*Copenhagen Business School, Computational Linguistics, ISV, Dalgas Have 15, DK-2000 Frederiksberg, Denmark, carl@iai.uni-sb.dk*

[2]*University of Tampere, Department of Computer and Information Sciences, Pinni B 1063, Kanslerinrinne 1, Tampere, Finland*

Eye-tracking is a growing field of research with several applications such as eye-movements in the recognition of natural environments, in human computer interaction, in speech and in reading. Eye-tracking has also been used as a means to understand the process of reading. This research focuses on the impact of single word properties, such as its duration, complexity (number of morphemes), familiarity of the word to the reader, its ambiguity, but also problems related to reading sequences of words in a context. Difficulties are usually measured by counting the duration of eye fixations on a particular word and the number of regressions to this word.

One of the aims within the Eye-to-IT project is to figure out what help a translator might need when reading a text with the goal to translate it into another language. It has been shown that reading - and thus Eye-fixation patterns - is very different in different reading purposes. Thus, [1] showed that reading a text of 200 words for understanding received 145 fixations, while there were almost 900 fixations when reading a similar text while typing a written translation of it. Besides the task of reading, fixation patterns also depend on the experience of the reader (or - for that purpose - on the experience of the translator), on the expected translation quality, on the difficulty and familiarity of the text to the reading subject, etc., etc.

Since many of the parameters causing particular fixation patterns are not well understood, we aim at creating a database of annotated texts, which can then be investigated in order to figure out the common patterns for different reading settings and to estimate the depending parameters.

The eye-tracking data and fixation patterns are obtained in an experimental setting, where a number of subjects are confronted with different reading tasks. An experiment supervisor would first introduce the subject to the task. The reading activities (and perhaps key-logging data, in case of written translation) would then be recorded by our software (Tobii 1750 eyetracker and Translog/GWM). In a third step the supervisor would go through the data together with the subject to assemble additional information.

This last step resembles a 'think aloud' protocol, with the difference that subjects see their own gaze activities on the screen and comment on previous actions. A subject would then explain where and what kinds of problem were faced during the reading task and what strategies he/she followed to solve those problems. The revision session also serves to readjust eye-mapping errors which occur when mapping the eye position data on a word. The tool, thus consists of the four following steps:

Collect **eye-position data,** i.e. the raw output data from eye-tracking devices. The TOBII 1750, for instance, produces a data sample every 20ms (at 50 Hz) consisting of the pixel positions (X,Y) for the left and the right eye and their pupil dilation

Eye-fixations are computed from the eye-position data and mapped on the characters and/or words looked at on the screen. This so-called **eye-mapping data** maps the eye-fixation data on (sequences of) symbols similar to the d*rift correction algorithms* [2]. Due to the inaccuracy of the technical devices and the visual system, this process is, however, not always correct. This is particularly so, if the reading task involves searching the screen where the gaze moves back and forth across different regions on the screen. A manual re-adjustment is therefore in many cases necessary.

During **mapping re-adjustment** the experiment supervisor and the subject go through the text that was read, together, and revise the eye-mapping data. Mapping re-adjustment is made possible with a tool in which the eye-position data is shown together with the eye-mapping data during a revision session. An example of a possible screen shot is shown in Figure 1. The dots show the eye-position data. Their size indicates the duration of the fixation, while the highlighted characters show the eye-mapping data. It is possible to review the reading session in real time, to slow it down, and to manually adjust erroneous eye-mapping positions. In addition, comments of the subject may be added, which reflect his/her mental state at a given moment.

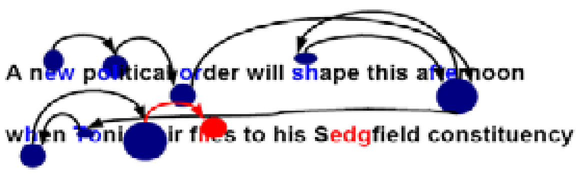


***Figure 1**: Sketch of the Annotation Tool.*

These **speak aloud annotations** will be stored with the mapping data in a separate field. In addition to 'free annotations', the subject will also be able to choose from a set of previously anticipated categories of mental states which best describe the subject's mental state at the given moment. The aim of this annotation is to enable statistical investigation of the annotated material in order to find typical patterns which lead to or entail the mental state.

The described annotation tool aims at 1. elaborating a standard for eye-movement annotation in the reading task, 2. making it possible to collect a substantial amount of comparable eye-movement data in order to 3. statistically investigate similarities and differences in eye-movement patterns across different reader, reading tasks and reading conditions and 4. to share our data, results and tools with other researcher.

References

1. Jakobsen, A. L. and K. T. H. Jensen (2007). Coordination of comprehension and text production in witten and oral translation tasks. Poster presented at the 13th Annual Conference on Architectures and Mechanisms for Language Processing (AMLaP) 2007.
2. Hyrskykari, A. (2006) Utilizing eye movements: Overcoming inaccuracy while tracking the focus of attention during reading. *Computers in Human Behavior* **22**(4), 657-671.

Web site quality evaluation combining eyetracking and physiologicial measures to self-reported emotions: an exploratory research

S. Prom Tep[1], A. Dufresne[2], .J. Saulnier[3], and M. Archambault[3]
[1]*RBC Financial Group Chair of E-commerce, HEC Montréal, Montreal (QC), Canada, Sandrine.prom-tep@hec.ca*
[2]*Department of Communications, Université de Montréal, Montreal (QC) Canada, Aude.dufresne@umontreal.ca*
[3]*User Experience Group, Web Bell Solutions, Montreal (QC), Canada, Julie.Saulnier@bell.ca, Marion.archambault@bell.ca*

Context

Studying Human Computer interaction, standard usability measures have focused on effectiveness, efficiency and satisfaction and do not incorporate emotions [1]. Recently, the importance of the affective quality in the design and in the evaluation of interactive systems has been stressed [2]. With e-commerce and the multiplication of retail Web sites, taking emotions into considerations become crucial as it has been clearly established that consumers are not always rational in their choices, nor utilitarian in their motivations [3]. If there is much more than cognitive information processing going on while consumers are online [4], measuring behavior for interactive systems evaluation has almost exclusively limited its focus to cognitive activities. Models combining emotions with human information processing variables are scarce and when available, they rely on reported data [5]. Research has shown though, that emotions are experienced both at the conscious and preconscious levels [6]. We will present an exploratory project, where we attempted to tackle this complex phenomenon. We triangulated non-consciously informed data derived from eyetracking and physiological measures, with post activity self-report of emotions on a scale to measure emotional reaction of users who are exploring a commercial Web site [7].

Objectives

Four working assumptions were made:

- The user profile (aesthetic and interest in the content) has an impact on how they will react to the imagery.
- There is a correlation between the user reported feelings and the variations in their physiological measurement during the interaction with the interface.
- Zones of interest in the interface can be detected using physiological measurement.
- The user's stress level in relation to the sensor diminishes throughout the session (habituation).

Methodology

Holbrook and Batra [7], present a set of 29 a priori multi-item emotional indices which categorizes emotions according to their type. While we did not seek to analyze all of these emotions through this experience, we selected some that are considered relevant in the field of human-machine interaction: joy (feeling pleased), surprise (feeling amazed), activation (feeling aroused, excited), hypoactivation (feeling bored), surgency (feeling playful), conflict (feeling tense), desire (feeling full of craving), involvement (feeling informed), déjà vu (feeling unimpressed), interest (feeling curious). We developed a questionnaire with these emotions and asked users to rate each emotion the interface provoked on a scale from 1 to 7 (e.g.: Did this interface make you feel bored? 1 meaning not at all bored and 7 meaning very bored). We also asked if they had felt any other type of emotion other than the ones proposed.

In order to explore the correlation between the self-reported emotional and behavioral reactions, we added physiological response to the usability testing of a commercial Web site, adapting the protocol to include physiological and eye-tracking measures of the user reactions. Among the range of sensors, we have chosen the least intrusive and the least dependent on the environment. So we used measures of respiration, skin conductance, temperature and pulse (BVP). Three scenarios were tested with six participants. In each session the participants were asked to verbalize their actions and reactions while completing the tasks. Each session was done in the same control environment and lasted ninety minutes. We wanted to measure their first impression as they saw the interface, so the experiment was organized with the following tasks:

- The users were greeted, received explanations and signed the agreement form.
- The sensors were installed and calibrated.
- Then for three sections of the site, the users were presented a page for 30 seconds, while physiological measures were taken; they were asked to fill the emotional self-report questionnaire, before moving to tasks related to that section of the site.
- We ended with a debriefing to get their impression on the experiment.

Testing environment

User testing was conducted in the Bell Solutions Web Laboratory. We used the MS Internet Explorer® Web browser (version 6.0) on a TOBII T120 unit (eye tracking system). We also used a 17" flat panel control monitor to observe the gaze movement. The processor was Intel® Core™2 CPU 6600@2.40Ghz 2GB of RAM, and the operating system was Windows® XP Pro2002 SP2. The software Tobii Studio 1.0 from Tobii was used to conduct the test and Biograph Infiniti from Thought Technology was used to capture the physiological measures. The testing room set-up conformed to the classic usability testing lab design with separate testing and observation rooms, with the control monitor connected to a computer through a VGA2USB adapter. We recorded comments from the user and the output of the control monitor, with an overlay of the gaze movement.

Analysis

Different analyses were conducted in relation to the hypotheses. First we compared self assessment measures with the different profiles of the subjects in terms of interest for a content and orientation toward aesthetics. Then we analyzed the evolution of the stress based on each physiological measure (i.e. increase in SC and pulse, temperature decrease, amplitude and frequency). We wanted to see whether stress was decreasing with habituation. We tested if we could find zones of interest where some of the physiological measures, especially those related to arousal, would be more affected. We also explored the evolution of interest as measured by the eyetracking system: pupil dilatation, fixation time, and correlated it with physiological measures of stress, globally and also in relation to specific zones of interest. Last but not least, we compared the measures trying to relate them to the self-assessment measures, correlating them with the global assessment, but also to each specific emotion, since they may differ in effect.

This is an exploratory research with limitations in the number of subjects and the length of the experiment, but the main objective was to develop methods and tools to combine the different view points from which to assess the quality of the user experience. One of our goals also seeked to explore the hypotheses and how they could be operationalized using the different measures we had in the context of the evaluation of a commercial Web site.

References

1. Norman, D. (2004). Emotional Design: Why we love (or hate) everyday things. New York. Basic Books.
2. Zhang, P., & Li, N. (2005). The importance of Affective Quality. *Communications of the ACM,* **48(9)**, 105-108.
3. Childers, T. L., Carr, C.L., Peckc, J., and Carson, S. (2001). Hedonic and utilitarian motivations for online retail shopping behavior. *Journal of Retailing,* **77(4)**, 511-535.
4. Ethier, J., Hadaya, P., Talbot, J., Cadieux, J**.** (2006). B2C web site quality and emotions during online episodes: An empirical study, *Information & Management,* **43**, 627–639.
5. Chamberlain, L. and Broderick, A.J. (2007). Qualitative Market Research: An International Journal **10(2)**, 199-216.
6. Hirschman, E. C., & Holbrook, M. B. (1982). Hedonic Consumption: Emerging Concepts, Methods and Propositions. *Journal of Marketing,* **46**, 92-101.
7. Holbrook, M.B., Batra, R. (1987). Assessing the Role of Emotions as Mediators of Consumer Responses to Advertising, *Journal of Consumer Research,* **14(3),** 404-419.

Measuring emotions using eye tracking

Jakob de Lemos, Golam Reza Sadeghnia, Íris Ólafsdóttir, and Ole Jensen
Research and Technology Department, iMotions - Emotion Technology A/S, Copenhagen, Denmark, technology@imotionsglobal.com

This abstract introduces a new automated method of measuring human emotions by analyzing eye properties via an eye tracking platform. The method utilizes eye tracking data in a new way that makes it possible to measure the immediate unconscious and uncontrollable emotional responses before they are cognitively perceived, interpreted, and biased by our mind.

Introduction

The eye property based method for quantifying emotions is implemented in Emotion Tool™ a software to measure immediate human emotional responses to visual stimuli. The non-intrusive, non-verbal, psychophysiology measurement instrument uses an eye tracking hardware and a statistical system to determine the level of excitement to an image and if the emotion is of a pleasant or unpleasant character. Each emotional response is extracted from subtle dynamical psycho-physiological changes within the respondents' eye via the eye tracking monitor. The results are hereafter classified and delivered directly as a fully quantified reading of the emotional response. Via the eye tracker the software collects and analyzes several subtle changes within the respondent's eye gaze characteristics, blink characteristics, and pupil change characteristics in order to determine the emotional response.

Accessing emotions through analysis of eye properties

The emotion measurement system (iET) can measure pleasantness and emotional reactions from humans looking at images, using a method that is nonintrusive, reliable and valid. Using specially designed algorithms and eye-tracking, we have invented a method and encapsulated it in a software which now enables researchers to quantify the basic low-level emotions, in an easy-to-use and time efficient way.

Psychophysiological measures

The iET system is an automation of the methods similar to the ones used by e.g. psychologists, witness experts and others, who have been trained to spot the subtle changes in facial expressions, to detect emotional signals from subjects. More specifically, the emotional signals are primarily based on the eye region. Pupil size, blink properties and gaze are analyzed and used to calculate an index of emotional reaction, using modern eye tracking equipment and specially developed software. Pupil size is known to be related to emotional reactions. For example, pupil dilation has been coupled with activation of the sympathetic nervous system [1-3]. However, the relationship is complex because pupil size is also related to cognitive processing load [4, 5] and the amount of light or hue in visual stimuli [1]. Blink has also been related to emotional reactions, for example with defensive reactions like emotion-modulated eye blink startle [6-8]. Finally, gaze patterns have been linked to emotional reactions [9].As an expert gains experience, his method becomes more accurate. One of these somatic markers mentioned above is not alone sufficient to spot an emotion in the subject. However several of these clues put together enable the experts to spot emotions correctly more than 80% of the time. In other words, the human brain is able to put this subtle information together in a way that can reveal the subjects' emotions. A similar technique as the one used by the experts has been formalized and is used within the iET emotion measurement system. This intuition that the expert has developed over a long time has now been formalized mathematically in an algorithm and software as a new easier way to gain access to the emotional reactions.

Conclusion

We propose an application area of eye tracking as nonintrusive, reliable, cost effective measurement of emotional response. This type of measurement is becoming increasingly more valuable for psychologists and emotion researchers to evaluate human behavior, memory and decision making – there is an obvious double benefit of both having respondents' visual attention combined with emotional response in one measurement. We demonstrate that this can in fact be done when stimuli are still images displayed on a monitor, in a controlled test environment.

References

1. Granholm, E. & Steinhauer, S.R. (2004): Pupillometric Measures of Cognitive and Emotional Processes. *International Journal of Psychophysiology*, **52**, 1–6.
2. Steinhauer, S. R., Boller, F., Zubin, J. & Pearlman, S. (1983): Pupillary dilation to emotional visual stimuli revisited, *Psychophysiology*, **20**.
3. Lang, P.J., Greenwald, M.K., Bradley, M.M. & Hamm, A.O. (1993): Looking at pictures: Affective, facial, visceral, and behavioural reactions. *Psychophysiology*, **30**, 261-27
4. Beatty, J. & Lucero-Wagoner, B. (2000): The pupillary system, in Caccioppo, J., Tassinary, L.G. & Berntson, G. (Eds.): *The Handbook of Psychophysiology*, Cambridge University Press, Hillsdale, New York
5. Staners, R.F., Coulter, M., Sweet, A.W. & Murphy, P. (1979): The papillary response as an indicator of arousal and cognition. *Motivation and Emotion*, **3(4)**, 319-340.
6. Bradley, M. M., Cuthbert, B. N. & Lang, P. J. (1999): Affect and the startle reflex, in Dawson, M.E., Schell, A. & Boehmelt, A. (Eds.): *Startle modification: Implications for neuroscience, cognitive science and clinical science*, Stanford University Press, Stanford, 242–276
7. Dichter, G.S., Tomarken, A.J. & Baucom, B.R. (2002): Startle modulation before, during and after exposure to emotional stimuli. *International Journal of Psychophysiology*, **43**, 191-196.
8. Ruiz-Padial, L., Sollers, J.J., Vila, J. & Thayer, J.F. (2003): The rhythm of the heart in the blink of an eye: Emotion-modulated startle magnitude covaries with heart rate variability. *Psychophysiology*, **40**, 306–313.
9. Calvo, M. G., & Lang, P. J. (2004). Gaze patterns when looking at emotional pictures: Motivationally biased attention. *Motivation and Emotion*, **28**, 221–243.
10. Tranel, D. (2000): Electrodermal Activity in Cognitive Neuroscience: Neuroanatomical and Neuropsychological Correlates, in Lane, R.D. & Nadel, L. (Eds.): *Cognitive Neuroscience of Emotion*,Oxford University Press, New York.

How epistemic motivations affect the perception and the visual exploration of advertising images

Stefano Livi, Marta Gentile, Sara Coletta, and Francesco Di Nocera
University of Rome "La Sapienza", stefano.livi@uniroma1.it

Introduction

According to Krugman [2], people differ to the extent they use different visual perception strategies (i.e. scanning and focusing) and such approach affect quality and quantity of images' memorization. From another point of view, social psychologists proposed that epistemic motivations are directly involved in environment information seeking. The present study's aim is to understand the role of an epistemic motivation, namely Need for Cognitive Closure (NCC), in the perception and exploration of advertising images. NCC was defined as a desire for a quick firm answer to a question, and could be a chronic tendency (i.e. individual difference) or could be activated by situational factors (e.g., noise or mental fatigue). The term "need" is here used to indicate a motivational tendency or an inclination that is function of an individual weight of costs and benefits of a closure (or a lack of closure) of the epistemic process [1]. From this point of view, we propose that epistemic motivation, directly affects both the perception of advertising images and the visual exploration strategy. By so doing it captures information scanning or focusing on images as a function of the individual epistemic motivational tendency and the ambiguity of the image.

Hypothesis

H.1 Relationship between NCC and fixation's number. This relationship should be negative, independently by the proposed stimulus because subjects that score high in the NCC scale grasp to the first information available, a "seizing" process aimed at reaching the decision quickly. In this way, this category of subjects immediately form an opinion. In parallel, observation style follows the same tactic, "focusing" with few and long fixations.

H.2 Relationship between number of fixations and message's perception of clearness. This relationship should be negative when subjects examine a low-creativity image because this kind of image needs a focusing process to be understood. On the other hand this relationship should be positive when a subject inspects a high-creativity image because this kind of image needs a scanning process with many short fixations.

H.3 Relationship between NCC and perception of clearness: this relationship should be positive when a subject that scores high in NCC scale examines a low creativity image, because of his/her intolerance for the ambiguity. On the other hand this relationship should be negative when the same type of subjects inspect a high creativity image.

Sample and Procedure

Thirty women, all Italian psychology students, aged between 20 and 30, participated to a single experimental session. About three months before the experiment, the dispositional measure was collected as well as the voluntary agreement to participate to a social psychology experiment. Nine advertising images were showed. In order to reveal and measure the exploration strategy of the subjects we analyzed their eye movements recorded while viewing two images, previously scored as the more and the less creative.

Measures

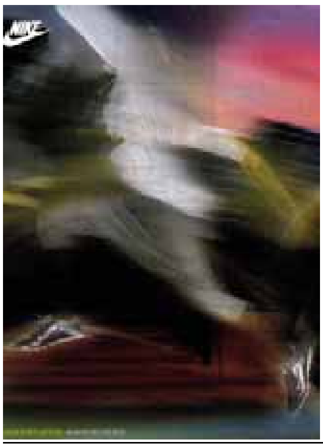
NCC: the shortened version of the NCC questionnaire [3] was administered.
Images clearness perception: Subject's perception of image's clearness and creativeness using two nine-step Likert's scale items.
Fixation's number and duration: An infrared eye-trackinger system was used to record the number and duration of fixations made by subjects while inspecting the images.

Results

Using the Pearson's coefficient we found that all of our hypotheses were verified.

Results showed that *high* NCC subjects perceived the image with *low* level of creativeness as having the message clearer; on the other side, *low* NCC subjects perceived the image with *high* level of creativeness as clearer. This distinction between

More Creative



Less Creative

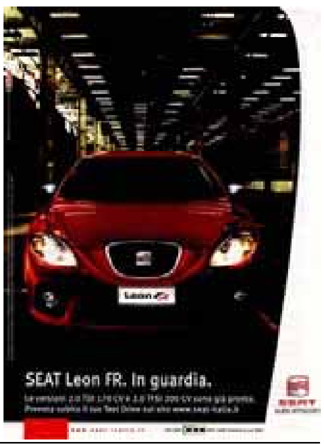
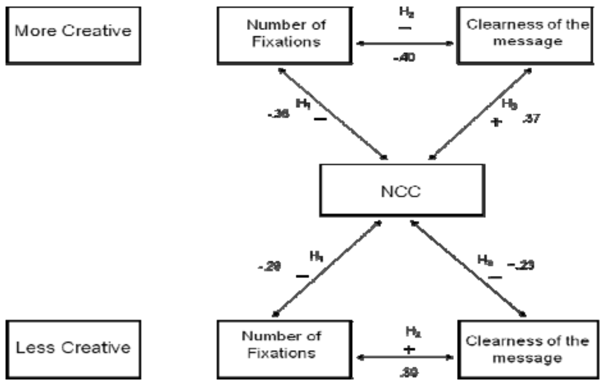


Figure 1: *Selected Stimuli.*

high and low NCC was related to different exploration patterns: in fact the results showed that *high* NCC subjects used a focusing process for inspecting (i.e. performed few and long lasting fixations), regardless of the proposed stimuli's characteristics. On the other hand, *low* NCC subjects inspected the images using a scanning process, performing many short fixations in the same time range.



General Discussion

Preliminary results indicate that epistemic motivations may guide the explorations strategy. Nevertheless, future research will be needed to investigate whether NCC drives the exploration strategy to some qualitative and semantic relevant elements of the each image. For instance, using eye-tracking data, it should be possible to define areas of interest that could disentangle more informative part that better fit the individual NCC tendency.

References

1. Kruglanski, W.A. (2004). *The psychology of closed mindedness Essays in Social Psychology*. Psychology Press New York.
2. Krugman, H. E. (1968). Processes underlying exposure to advertising. *American Psychologist* **23(4)**, 245-253.
3. Pierro, A. & Kruglanski, W.A. (2004). Revised Need for Cognitive Closure Scale. *Submitted*.

Analysis of Advertising Effectiveness with Eye Tracking

G. Theuner[1], K. Pischke[2], and T. Bley[2]
[1]Department of Marketing, Ludwigshafen University of Applied Science, Ludwigshafen, Germany, theuner@fh-lu.de
[2]UsabilityAgent, Dresden, Germany, info@usabilityagent.de

Introduction

Global business organizations as well as medium and small sized companies use increasingly different kinds of outdoor and indoor advertising like posters and billboards. Analysis of advertising effectiveness has become a more and more important area of corporate communication. This also results from the fact that the existing methods have not proved to be as effective as anticipated. Eye Tracking for example is a well proven instrument to measure the perception of a print advertisement. But it cannot be used outdoor for motorists and cyclists.

Poster and billboards in particular have to put their message across to pedestrians in a few seconds, and for motorists and cyclists in a split second. [1]. Hence the design of the message is very compact, mostly in the form of pictures and a short claim. But often the question arises if the important elements and information will be noticed in the real everyday situation.

The main objective of the current research project was to develop an instrument what allows to measure the participants' perception and recall level in a virtually real traffic situation. This applies particularly to target groups like car drivers and pedestrians.

Experimental Design

The initial point of the research project was to create a standardizing body for an easily applicable, cost-effective pre and post test to determine the success of outdoor and indoor advertisement.

In the first step the participants are provided with a computer animated real traffic situation with advertisement on the left or right side of the street. The driver's speed is variable as well the poster/billboard position. This test instrument is called AdSpectator. The test is covered by a survey:

- pre questionnaire about attitude, interest and knowledge concerning the topic and the company

- post questionnaire (depends on advertiser objectives) about recognition and unaided/aided recall relating to the different design elements and to the message. (Figure 1) The results are based on a questioning only (without observation). One advantage of the single test is the possibility for testing a group with one notebook and projector only.

In a second step eye tracking was combined with the so called AdSpectator-analysis (Figure 1). The combined test provides subjective results from the single test, and objective data about the actual perception per single participant.

Eye tracking is a method used to verify where a person is looking. The eye tracker provides objective and quantitative evidence of the user's visual process. Eye movements are generally recorded to determine a user's attention patterns concerning a given stimulus [2]. Eye tracking cameras can:

• Tell whether users are looking at the poster/billboard. Without an eye tracker, it is impossible to verify exactly where users are looking.

• Tell whether users are reading or scanning. It is easy to distinguish reading, a user's systematic fixation on word clusters, from scanning for particular words or phrases.

• Tell where the first view is directed at the poster [3, 4].

• Measure the relative intensity of a user's attention to different parts of a poster/billboard. By dividing the projection screen into areas of interest (AOI), such as the message or pictures, one can see when and how long a participant looks at each AOI.

• Compare scan patterns of all participants. By counting how long each participant looked at each area, and in what order, scan patterns of groups of participants can be compared.

It is important to remember that an eye tracking device can give an insight of a participant's actual perception, but only in conjunction with an adjacent questioning of participants. Just because the eye looks at something does not automatically mean that what it sees is consciously perceived by the brain. It is therefore necessary to question the participants in order to find out what the users remember seeing.

The software **AdSpector** is a new software for testing poster advertisement „in use". AdSpector uses 3D graphics to simulate a car ride or a walk through a row of houses or inside a room. The poster which should be evaluated appears

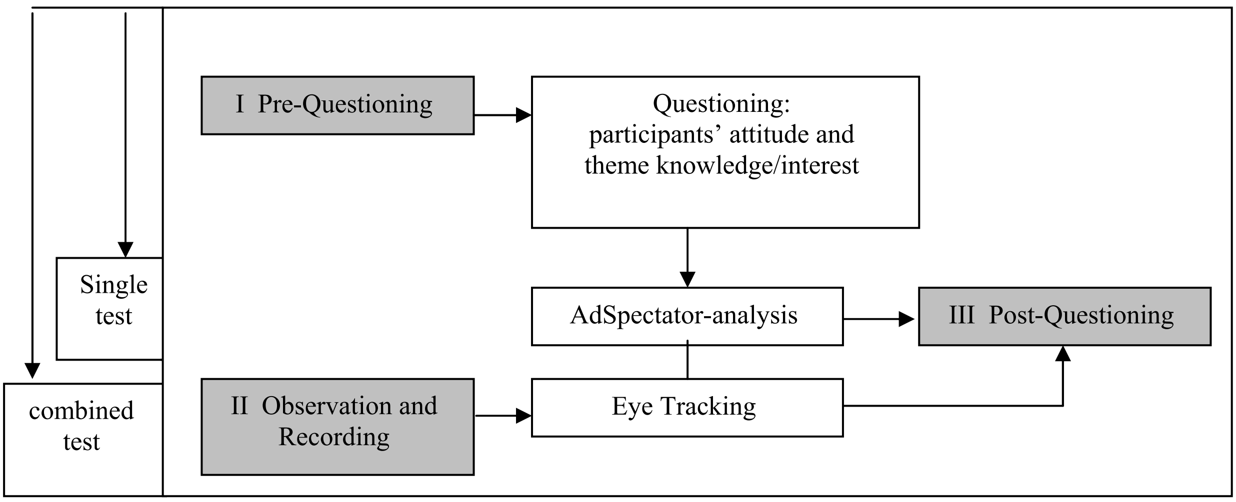


***Figure 1**. Experiment Procedure*

on a billboard, a city light or an advertisement pillar which are integrated in the real scene. So by the help of AdSpector the test persons drive or walk past the poster in real time (projector screen). Afterwards they are asked which information they have seen, how they were emotionally involved. AdSpector simulates the fleeting period of observation which is typical for poster ads.

Figure 2*. AdSpector combined with the Eye Tracking Path*

Research Outlook

First tests have shown that the data deliver fundamental conclusions as how to improve the effectiveness of advertising. AdSpector application and the whole tests are flexible. They can easily be extended and adapted to special environmental conditions, company needs and objectives, and in every region. Advertising effectiveness stated by the samples of this test will be fine tuned by collecting and analyzing data on the following tests. Tthe new samples will be involved to a larger extent in the interpretation of the results as this will reinforce the recommendations arrived at.

References

1. K.Rayner (1998). Eye Movements in Reading and Information Processing – 20 Years of Research, In: *Psychological Bulletin*, Vol. **124 (3)**, , p. 398.
2. A.T. Duchowski (2003). Eye Tracking Methodology, Theory and Practice, Wiesbaden, p. 20-26
3. G.Bente (2004). Recording and Analysis of Gaze behavior, In: *Textbook Mediapsychology*, Mangold, Vorderer, Bente (Editors), Göttingen, p. 319.
4. R.Grier, P. Kortum, J. T. Miller (2006). How Users View Web Pages – An Exploration of Cognitive and Perceptual Mechanisms, pp. 22-41, http://aptima.com/publications/2006-InPress_Grier_Kortum_Miller.pdf (07/2007)

Multivariate analyses for the study of behavior: an integrated approach

Maurizio Casarrubea
Università degli Studi di Palermo, Facoltà di Medicina e Chirurgia, Dipartimento di Medicina Sperimentale, Sezione di Fisiologia Umana, Palermo, Italy, m.casarrubea@unipa.it

Introduction

Simple quantitative evaluations of isolate behavioral elements (i.e. frequencies, durations, per cent distributions) are not representative of the whole behavioral structure [1]. As suggested in a landmark paper from Spruijt and Gispen [2], it is only through the evaluation of the inter-relations among behavioral elements that it is possible to explore behavior from very different points of view, greatly beyond what the human eye can intuitively interpret. In the present paper a brief outline of different multivariate techniques for behavioral analyses will be presented in the attempt to underline the feasibility of their integration.

Approaching transition matrices

Different levels of multivariate analyses (MVA) can be described on the basis of computational requirements and complexity of the approach. The first step is the construction of an ethogram that is a formal list containing descriptions of behavioral elements. After that, using specific software coders such as The Observer (Noldus Information Technology), behavioral elements have to be coded from the collected video-files and transitions from an element to another one traced in a transition matrix (TM). In brief, a TM is a table representing shifts among the behavioral elements, according to the selected ethogram [3][4][5][6]. A first and relatively simple approach to the analysis of a TM can be represented by the so called stochastic analysis [4][5]. A stochastic analysis requires transition matrix to be transformed into a matrix containing relative frequencies of transition from a given behavioral element to the others. Matrices containing relative frequencies can be graphically expressed through pathway diagrams. Figure 1 illustrates an example of pathway diagram representing *probabilistic relations* among five different behavioral elements identified in a group of 42 rats observed in open field. Three probability ranges were selected: between 0.10 and 0.24 (thin dotted arrows), between 0.25 and 0.49 (medium arrows), and between 0.50 and 1 (large arrows).

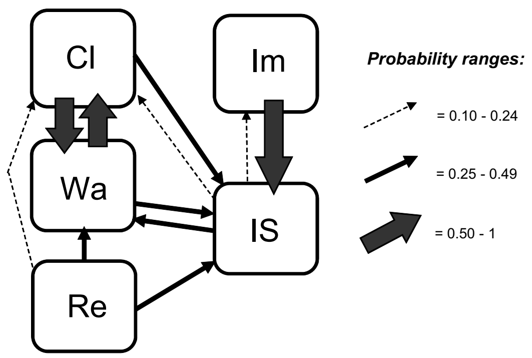


***Figure 1**. Transitional probabilities among 5 different behavioral elements observed in 42 rats in open field. Different probabilities of transition are depicted by arrows of different thickness. Cl = Climbing, Im = Immobility, Wa = Walking, IS = Immobile-Sniffing, Re = Rearing.*

A higher level of approaching matrices is represented by the cluster analysis which transforms transition matrices into similarity tables through an aggregative procedure. Cluster analysis allows the identification of behavioral clusters. The result of such procedure is a matrix containing only absolute affinity values, i.e. an half matrix where each cell indicates the "vicinity" between two given elements. Cluster analysis could be considered in some extent less intuitive than the stochastic one because of the underlying aggregative algorithm that converts transitions into similarity values (i.e. the direction of the behavioural flow is not expressed). Main outcome of cluster analysis is a dendrogram that is a tree diagram. Dendrogram in figure 2 represents *"vicinity" relations* among the same five behavioral elements presented in figure 1. However, even if both stochastic and tree diagrams present behavior in a graphically intuitive way, underlying matrices may need further statistical analyses.

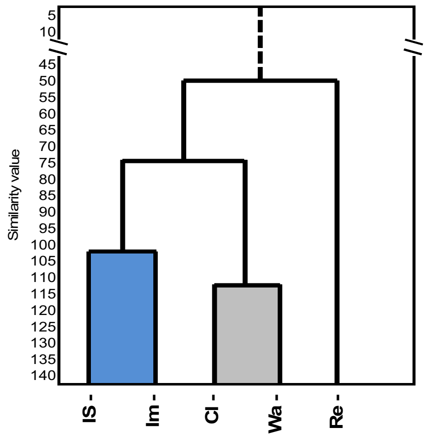


***Figure 2**. Dendrogram representing similarities among 5 different behavioral elements observed in 42 rats in open field. IS-Im and Cl-Wa are closely linked each other molding two different clusters. For abbreviations see figure 1.*

Transitional matrices statistics: adjusted residuals

To compare transitional matrices, an useful method has been suggested by Tavaré et al. which proposed a correction factor – the so called gamma correction – to multiply by the chi-square value [7]. However, an elegant method to assess significance of cells within matrices is the one used by B.S. Everitt [8] and B.M. Spruijt [2] and, after them, by different Authors [5] [9]. Through this method a transition matrix is transformed into a matrix containing adjusted residuals. In summary, for each given transition, adjusted residual represents the difference between the observed cell value and an expected one, the latter calculated on the basis of a random distribution of transitions. Positive residuals indicate transitions occurring more often than expected, negative residuals represent transitions occurring less often than expected. Adjusted residuals can be evaluated through MatMan 1.1 (Noldus Information Technology), i.e. a specific software for matrix manipulation and analysis. The consistent advantage of adjusted residuals is that they can be expressed according to a Z-distribution so that P-values can be easily found in a common Z-table. Advantages provided by dendrograms and stochastic pathway diagrams is that they represent patterning among behavioral elements in a simple and intuitive way. However, both stochastic and cluster analyses are extremely sensitive to noise (i.e. disturbing and/or uncommon behavioral elements need to be removed before running the analysis), moreover they are close to "snapshots" of the whole observational period.

Events along time: T-patterns

T-pattern detection doesn't require an a priori "noise reduction" and, more important, it represents events *along* the session time. T-pattern analysis is carried out through the specific software Theme (Patternvision and Noldus Information Technology). This program performs a recursively test, checking the distribution of every combination of events *along* a specific time window [1].

Figure 3 represents an hypothetical T-pattern of four elements. If the time lag of an events sequence is not randomly distributed a simple T-pattern is detected (elements "e" and "d" in figure 3). In following steps, above simple t-pattern is processed again and if there is a temporal relation with other events, they are combined into higher order T-patterns (e-d-l-n, figure 3), and so on, repetitively, following a "bottom-up" process [1]. Since the graphical representation of T-patterns and the results of cluster analysis are both visualized through tree diagrams (figure 2 and figure 3), it is important to mention that cluster analysis is based on the similarities between events (figure 3). On the other hand, the tree structure provided by Theme does not represent such similarities, but the existence of significant relationships along time. Figure 4 shows an highly recurring T-pattern found in ten subjects randomly taken from the main group of animals represented in figure 1 and 2.

Discussion

An even swift comparison among figures 1, 2 and 4 makes clear how these three, rather different, multivariate approaches strengthen each other in representing animal behavior. What's more, each representation perfectly fits with the remaining ones: pathway diagrams and dendrograms show patternings among behavioral elements from stochastic and "aggregative" points of view respectively. On the other hand, T-patterns represent the behavioral structure along time. The results presented here show that these different MV approaches can be successfully used together for a better and more realistic description of behavioral patternings (i.e. to identify the presence of possible relationships between the elements of a behavioral sequence).

It is my contention that even hundreds of purely descriptive parameters make available only a partial and/or incomplete view of a given behavior. In other words, descriptive analyses reduce behavior to simple numbers. This is a reductionistic conception similar to the so called Cartesian Reductionism [10]. This approach does not work for behavioral analyses because a behavior is characterized by emergent phenomena arising from the inter-relations among events. Thus, quantitative evaluations such as durations, per cent distributions or frequencies of disjointed behavioral elements are not able to offer answers to *the* crucial question in all behavioral studies: what about the relationships between the observed elements? It is exactly here that descriptive approaches to behavioral study should give way to the multivariate analyses.

Acknowledgements

The author would like to thank Prof. G. Crescimanno, Prof. F. Sorbera (Università degli Studi di Palermo, Italy) and Prof. F. A. Lopes (Universidade Federal do Rio Grande do Norte, Brazil) for their kind and constructive suggestions provided during the preparation of the present paper.

References

1. Magnusson, M.S. (2004). Repeated Patterns in Behavior and Other Biological Phenomena. Eds: Oller, D.K, Griebel, U. *Evolution of Communication Systems: A Comparative Approach*. Cambridge: MIT Press 111-128.
2. Spruijt, B.M., Gispen, W.H. (1984). Behavioral sequences as an easily quantifiable parameter in experimental studies. *Physiology & Behavior* **32**, 707-710.
3. Hemerik, L, Bukovinszky, T., Gols, R., van Lenteren, JC., Vet, L.E.M. (2006). Enter the matrix: How to analyze the structure of behavior. *Behavioral Research Methods* **38**, 357-363
4. Casarrubea, M., Sorbera, F., Crescimanno, G. (2006). Effects of 7-OH-DPAT and U 99194 on the behavioral response to hot plate test, in rats. *Physiology & Behavior* **89**, 552-562
5. Casarrubea, M., Sorbera, F., Crescimanno, G. (2007). Multivariate analysis of the modifications induced by an environmental acoustic cue on rat exploratory behavior, *Physiology & Behavior* doi: 10.1016/*j.physbeh*.2007.11.010
6. Espejo, E.F., Mir D. (1993) Structure of the rat's behaviour in the hot plate test. *Behavioral Brain Research* **56**, 171-176.
7. Tavaré, S., Altham, P.M.E. (1983). Serial dependence of observations leading to contingency tables, with corrections to chi-square statisticss. *Biometrika* **70**, 139-144
8. Everitt, B.S. (1977). *The analysis of contingency tables*. New York: John Wiley and Sons
9. van Lier, H., Coenen, A.M., Drinkenburg, W.H. (2003). Behavioral transitions modulate hippocampal electroencephalogram correlates of open field behavior in the rat: support for a sensorimotor function of hippocampal rhythmical synchronous activity. *The Journal of Neuroscience* **23**, 2459-2465
10. Mikulecky, D.C. (2001) The emergence of complexity: science coming of age or science growing old? Computers and Chemistry **25**, 341-348.

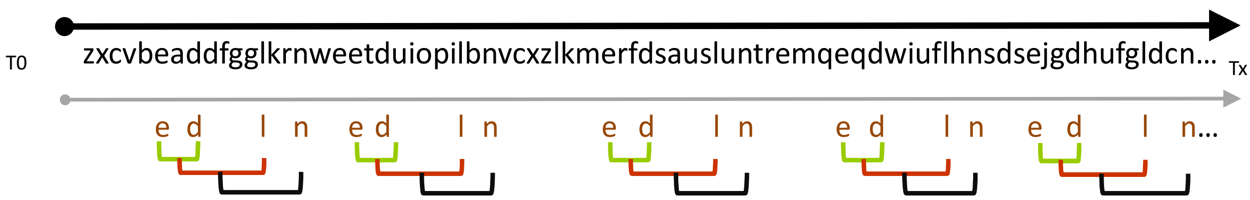


Figure 3. *Black upper arrow: observational period ($T_0 - T_x$) consisting of 26 hypothetical behavioral events (letters). The T-pattern represented in the bottom grey arrow (e-d-l-n) becomes evident when all the others behavioral occurrences are left out from the whole observational time line (black upper arrow).*

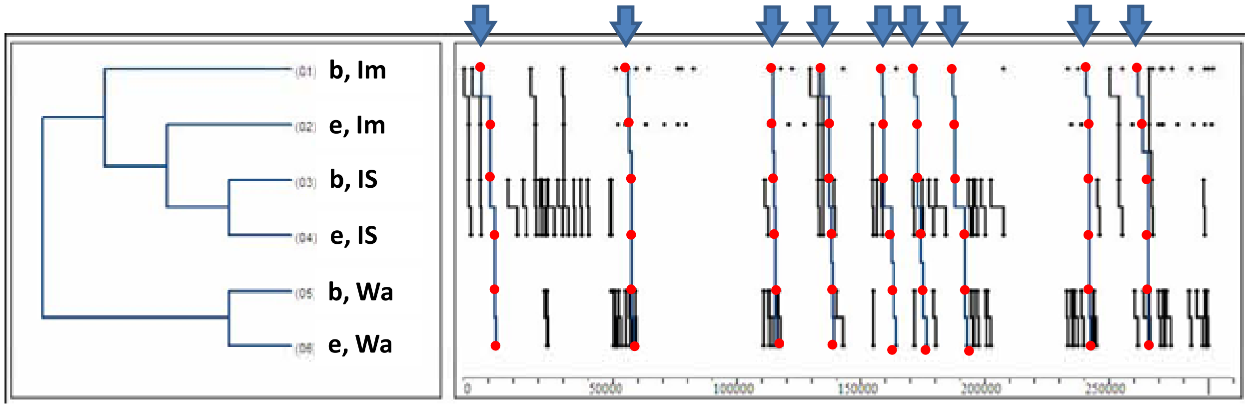


Figure 4. *T-Pattern encompassing Immobility, Immobile-Sniffing and Walking. T-pattern structure is shown in the left panel while the occurrences of the responses (dots) along x-axis (from frame 0 to frame 30000) are shown in the right panel. Larger dots = events encompassed in the T-pattern. b = beginning, e = ending. For other abbreviations see figure 1.*

Motion analysis of butterflies: A method to investigate the evolutionary origin of the unique pollen feeding behaviour of *Heliconius* butterflies (*Nymphalidae*)

A.L. Hikl, H.W. Krenn, and S.H. Eberhard

Department of Evolutionary Biology, University of Vienna, Vienna, Austria, antschal@gmail.com

Pollen feeding behaviour is unique to butterflies of the neotropical genera *Heliconius* and *Laparus* (Heliconiinae, Nymphalidae) and may be key for understanding life-history evolution [1]. Like many other butterflies they feed on floral nectar but they actively collect pollen on their proboscises and use it as an additional food source. Amino acids are extracted from the pollen grains by way of an extra oral digestion involving saliva [2] [3]. The mechanisms and the evolutionary origin of this special behaviour are unknown.

Here, we address the idea that pollen feeding behaviour may have originated from proboscis cleaning behaviour by using motion analysis. We compared mouthpart movements of 9 butterfly species (6 individuals each) of different relatedness [4]: 5 pollen feeders of genus *Heliconius* (*H. cydno, H. hecale, H. melpomene, H. pachinus* and *H. sara*) as well as 2 closely related species, pollen feeding *Laparus doris* and non pollen feeding *Eueides isabella*; *Dryas julia*, a related non pollen feeding Heliconiinae; and non pollen feeding species *Anartia fatima* as nymphalid out-group. All species were collected in the field near the "Tropical Station La Gamba" in Costa Rica and set free after use in experiments. To initiate pollen extraction behaviour or proboscis cleaning behaviour pollen or small glass beads (ca. 106µm) were placed on the proboscises of all butterflies. The behaviour of butterflies was recorded on video tape with a JVC GZ-MG37E hard disc camcorder for 20 min each. Three distinct proboscis movements, 5 different degrees of proboscis extensions, movements of the entire butterfly and the release of saliva were coded from video tape by using "The Observer XT" [5] [6]. All movements and the release of saliva were coded as state events, with a start point and an end, for comparing the quantity and the duration between species.

Preliminary results show that similar patterns of movements are employed by pollen feeding and non pollen feeding butterflies suggesting that the evolutionary origin of pollen feeding behaviour is likely a modification of proboscis cleaning behaviour. Furthermore there are also indications that the key innovations of pollen feeding behaviour are the frequency of the proboscis movements and the repeated release of saliva during pollen extraction. Motion analyses using "The Observer" software make it possible to investigate the pollen feeding behaviour of *Heliconius* butterflies on a fine scale and therefore new insights in the evolution of this unique behaviour can be achieved.

The study was funded by the Austrian Science Fund (FWF Project 18425-B03).

References

1. Gilbert, L. E. (1972). Pollen Feeding and Reproductive Biology of *Heliconius* Butterflies. *Proc. Nar. Acad. Sci.* **69(6)**, 1403-1407.
2. Krenn et al., in progress.
3. Eberhard S. H., Hrassnigg N., Crailsheim K., Krenn H. W. (2007). Evidence of protease in the saliva of the butterfly *Heliconius melpomene* (L.) (Nymphalidae, Lepidoptera). *Journal of Insect Physiology*. **53**, 126-131.
4. Penz CM. (1999). Higher level phylogeny for the passion-vine butterflies (Nymphalidae, Heliconiinae) based on early stage and adult morphology. *Zoological Journal of the Linnean Society* **127**, 277-344.
5. Noldus Information Technology (2005). The Observer XT (software).
6. Noldus, L.P.J.J., (1991). The Observer: a software system for collection and analysis of observational data. *Behavior Research Methods, Instruments, and Computers* **23**, 415–429.

Biting behavior induced by acute stress in the rat during experimental tooth movement

J.L. Zeredo[1], Y. Kumei[3], T. Shibazaki[2], N. Yoshida[2], M. Kimoto[1,4], and K. Toda[1]

[1] Division of Integrative Sensory Physiology, jllzeredo@gmail.com, and [2] Division of Orthodontics and Dentofacial Orthopedics, Graduate School of Biomedical Sciences, Nagasaki University, Nagasaki

[3] Section of Biochemistry, Department of Hard Tissue Engineering, Graduate School, Tokyo Medical and Dental University, Tokyo,

[4] Physiological Laboratory, Japan Women's University, Tokyo, Japan.

Biting behavior and emotional stress are directly related in rats [1]. In humans, likewise, emotional stress is known to induce parafunctional oral behaviors such as bruxism, tooth clenching, and nail biting. This phenomenon is well documented, but its mechanism is poorly understood. Would there be a purpose for the development of these stereotypies? One explanation is that such behaviors would be sought in order to counteract environmental stimuli. Indeed, biting is known to help alleviate the stress response in rats [2-4]. However, it is not clear what role the sensory experience from tooth contact plays in emotional experience in general and stress in particular. For example, the response to pain, a sensation with strong emotional component, includes not only stimulus avoidance, but also behaviors that actively inhibit pain, such as rubbing or licking the affected area. It is conceivable that stressed animals would seek the pleasurable or familiar sensation of tooth contact as a means to reduce stress. To address this question, we measured the biting behavior in rats during mechanical tooth movement and compared to that of naïve controls. Previous studies have shown that periodontal sensation is reduced during mechanical tooth movement [5].

Under barbiturate anesthesia, Wistar albino rats weighing 150 g had orthodontic springs placed on their maxillary incisors. The springs were made of Co-Cr wire with 0.4 mm in diameter and delivered a force of about 40 g in the lateral direction. In the same procedure, electromyography (EMG) electrodes were implanted into the masseter muscle on one side. Control rats went through a sham procedure, and also received EMG electrodes. EMG electrodes consisted of two stainless steel wires insulated except for 3 mm at the tip. One end of the wires was passed through the muscle 3 mm away from each other; the other end was passed from under the skin toward the top of the head and soldered to a connector, which was fixed to the skull with dental cement. After 24 h, the rats' response to acute restraint stress was recorded. The rats were placed inside a restraining cylinder with an opening at the snout for 30 minutes, and after that, biting behavior was measured by the masseter muscle EMG and by video monitoring of the freely-moving animals for 5 minutes. During recordings the animal was connected to the EMG apparatus by a flexible cable attached to its head. The cable was kept elevated to avoid entanglement without restraining animal movement. The EMG signal was passed through a differential amplifier and stored in a computer. Biting activity was estimated on video by observing the rapid bulging movement of the eyes and the movements of the cheeks and whiskers. The duration of these movements was measured with a stopwatch. EMG activity was analyzed by counting the number of bursts in the recorded waveform.

As a result, after restraint stress biting behavior was similar between tooth movement and control rats, both in duration and number of strokes. Biting behavior is a reliable parameter of stress response in rats. The data indicate that the sensory inputs from the teeth are not essential to the expression of stress-induced biting behavior. It is suggested that biting behavior is a stereotypical component of the stress response, but not necessarily a behavior sought with the purpose of relieving stress through a comforting sensory experience.

This study was supported by a grant from the Japan Society for the Promotion of Science to JLZ (No. 19791589, FY 2007-2008)

References:

1. Rosales, V.P., Ikeda, K., Hizaki, K., Naruo, T., Nozoe, S., Ito, G. (2002) Emotional stress and brux-like activity of the masseter muscle in rats. *European journal of orthodontics* **24**: 107-117.
2. Okada, S., Hori, N., Kimoto, K., Onozuka, M., Sato, S., Sasaguri, K. (2007) Effects of biting on elevation of blood pressure and other physiological responses to stress in rats: biting may reduce allostatic load. *Brain research* **1185**: 189-194.
3. Tsuda, A., Tanaka, M., Ida, Y., Shirao, I., Gondoh, Y., Oguchi, M., Yoshida, M. (1988) Expression of aggression attenuates stress-induced increases in rat brain noradrenaline turnover. *Brain research* **474**: 174-180.
4. Vincent, G.P., Pare, W.P., Prenatt, J.E., Glavin, G.B. (1984) Aggression, body temperature, and stress ulcer. *Physiology & behavior* **32**: 265-268.
5. Nakanishi, H., Seki, Y., Kohno, T., Muramoto, T., Toda, K., Soma, K. (2004) Changes in response properties of periodontal mechanoreceptors after experimental orthodontic tooth movement in rats. *The Angle orthodontist* **74**: 93-99.

Ethologic observation of rats and mice as a support for the strengthening of the Russell's "three Rs"

Jorge Enrique Bueno Prieto
Department of Biology, Universidad Nacional de Colombia, Biotechnology Institute, Biomimetics Laboratory, jebuenop@unal.edu.co

Abstract

The use of animals in scientific experimentation causes a dual reaction between those who defend the use of non-human animals to certify practices and experimental processes, and those who are against of the use of animals based on the ethical principle of respect to other forms of life. The implementation of alternatives to the use of animals for experimentation has had its equilibrium point regarding the two former fronts, in the Russell's Rs (Reduction, Refinement and Replacement) [2]. The behavior study has became the main tool to transform the animal-world vision, especially of those animals used for scientific experimentation, which were before seen as machines; and in a slow but firm manner, such a conception has been turned into ethical principles for the animal handling in research [5]. Rats and mice are currently the most used animals in the scientific search. There are behavioral parameters known of wild animals, but in dealing with specific strains, Latin America's studies area scarce. This is the reason why an ethological evaluation program on the species *Rattus Novergicus* (Wistar) *and Mus Musculus* (BALB /c – C57BL/6) [4], spf animals *(Specific Pathogens Free)* of the laboratory of biological reagents of Biotechnology Institute from Universidad Nacional de Colombia was started.

It has been chosen the video record system by a laptop computer and a web camera because it offers a reliable and cost-effective follow up, and in addition it is a non distraction element for the animals under observation. Both the rats and mice are placed in a micro-environment (box one cage) with permanent food and drink [1], and by the time of being submitted under observation, the video system is placed at a distance of approximately 2 meters away. Such a system allows that the collected data are based on an activity in place without any human presence and for the time esteemed by the researcher.

The collected information is unified in ethograms, and there the data are discriminated in behavioral patterns, specific units of both individual and group action, adaptability conditions and a brief description of the observed actions [3]. Finally, action-time frequencies of the behavioral units are established. Up to now low aggressive levels in mice and non-existent levels in rats are found, also found are the high preference in the use of enrichment tools and high interaction between individuals, as well as the establishment of activity "times" for feeding, nattiness and rest, with scarce modification under alteration external conditions, which is a sign of high adaptability. With this information, patterns of improvement and attention are generated for the animals' well-being in our laboratory and so to guarantee a responsible and ethical management, so as to strengthen the Russell-Burch's Reduction, Refinement and Replacement principles.

References

1. Booth, D.A. (1972). Conditioned satiety in the rat. *J. Comp. Physiol. Psychol.* **81**, 457–71.
2. Cardozo de Martínez, C.A., et.al. (2007). *El animal como sujeto experimental. Aspectos técnicos y éticos*. [Santiago de Chile]: Universidad de Chile, Centro Interdisciplinario de Estudios en Bioética, 2007.
3. Crawley J.N., Paylor R. (1997). A proposed test battery and constellations of specific behavior paradigms to investigate the behavioral phenotypes of transgenic and knockout mice. *Horm. Beha.* **31**, 197–211.
4. Self, D.W. (1999). Comparison of transgenic strategies for behavioral neuroscience studies in rodents. *Psychopharmacology* **147**, 35–7.
5. Spruijt, B.M.; Hol, Th.; Rousseau, J.B.I. (1992). Approach, avoidance and contact behavior of individually recognized animals automatically quantified with an imaging technique. *Physiology & Behavior* **51**, 747-752.

Effects of metabotropic glutamate receptor 5 antagonist MPEP on anxiety-like behavior in immature rats

Anna Mikulecká and Pavel Mareš

Institute of Physiology, Academy of Science of the Czech Republic, Prague, Czech Republic, nmikul@biomed.cas.cz

mGluR5 subtype of metabotropic glutamate receptors is present in various brain regions such as hippocampus, amygdala and cortex. An anxiolytic-like activity of the selective mGluR5 antagonist MPEP has been demonstrated in several conditioned and non-conditioned tests of anxiety in adult animals (for a review see [2]). In our laboratory, MPEP exhibited anticonvulsant actions in both adult and immature rats without inducing side effects on motor performance [3,4]. There are no studies dealing with effect of MPEP on behavior in developing animals. Therefore, we designed the present study to examine possible anxiolytic-like effect of the drug in immature rats particularly within the anticonvulsant dose range.

The experiments were performed in 18-, and 25-day-old Wistar rats. Experimental animals received i.p. 10, 20 or 40mg/kg of the drug, controls were injected with saline. Behavioral testing was performed 15 min (1st session) and was repeated 60 min (2nd session) following drug administration in the light-dark box [1]. The light-dark test is based on the tendency of animals to explore a novel environment vs. the tendency to avoid the aversive properties of a brightly lit arena. The individual rat was placed into the black part of the box facing the light part and its behavior was registered on a videotape for 5 min and scored using the Observer (Noldus Information Technology). The following parameters were evaluated: the time spent in light part, the number of transition between both parts and the number of squares crossed in the light part.

Compared with the controls, all doses of MPEP increased the time spent in the light part (a measure of anxiety-like behavior) in the 1st session in both 18- and 25-day-old rats, the effect of 20 mg/kg dose being more expressed. All MPEP doses also increased the number of transitions between light and dark parts as well as the number of squares crossed in the light part in the 1st session in either age group. There was a marked decrease in all measured behavioral parameters in the 2nd session.

In conclusion, our data suggests that MPEP exerts anxiolytic-like activity also in immature rats without any sedation or suppression of locomotion. The decreased behavioral response in the 2nd session show that prior experience with the light-dark box may alter subsequent behavioral responsiveness of the immature animals to the light-dark test of anxiety.

This study was supported by a grant Nr. 305/06/1188 of the Grant Agency of the Czech Republic.

References

1. Costall, B., Jones, B.J., Kelly, M.E., Naylor, R.J. and Tomkins, D.M. (1989). Exploration of mice in a black and white test box: validation as a model of anxiety, *Pharmacol Biochem Behav* **32**, 777-85.
2. Lea, P.M.t. and Faden, A.I. (2006). Metabotropic glutamate receptor subtype 5 antagonists MPEP and MTEP, *CNS Drug Rev* **12**, 149-66.
3. Lojková, D. and Mareš, P. (2005). Anticonvulsant action of an antagonist of metabotropic glutamate receptors mGluR5 MPEP in immature rats, *Neuropharmacology* **49**, Suppl 1, 219-29.
4. Mareš, P. and Mikulecká, A. (2004). MPEP, an antagonist of metabotropic glutamate receptors, exhibits anticonvulsant action in immature rats without a serious impairment of motor performance, *Epilepsy Res* **60**, 17-26.

Hierarchy in primates: combining observational and agent-based simulation methods

Ignasi Cifre and Francesc Salvador i Beltran
University of Barcelona, Barcelona, Spain, icifre@ub.edu

Hierarchical structure is a fundamental trait of primate societies and largely conditions the behavior of individuals living together in a group. Which laws are implied in the formation and structure of these hierarchies? Why do the most dominant subjects in certain groups stand in the center of the group, whereas in other groups of the same species they do not? In the field of animal behavior research, many different hypotheses have been proposed, but none has been entirely satisfactory, given the many variables in the set. A theory that fully explains all the empirically observed complexity has not been developed. Based on the adaptive-behavior modeling approach, we have used computer programming to implement many of the variables that authors such as [1] and [2] consider important when structuring hierarchical formations in primates. Both the type of food (durability, size, etc.) and the consequences of being at a specific place, with specific space and food restrictions, can be two of the most important variables determining which kind of hierarchy structure will be formed by a group. [1] relates the type of distribution of the resources to different strategies of acquisition of these: a) the *within group scramble* (WGS) determined essentially by a type of food that is spatially dispersed and therefore can not be monopolized, b) the *within group competition* (WGC) given fundamentally with resources very spatially clumped so food can be monopolized. Following the model proposed by [3], we have programmed an agent-based simulator called DOMLogo in which the rules of behavior prescribed allow complex hierarchical structures to emerge. DOMLogo (see figure 1) was implemented with multi-agent simulation platform Starlogo 2.1 [4]. This platform was created for simulating non-centralized systems of agents. The main feature of these systems lies in that there is no leader regulating or controlling the behavior of the other agents; each agent in these groups is totally autonomous in its behaviors at every moment. Starlogo allows us to program agents behavior in parallel (where all the agents in the simulation act at the same time), a feature that is very important when we try to simulate social interactions, since in nature they occur in a parallel but not in a serial way (acting in turns) as done in other software not created specifically for the simulation of agents. In a first phase of this research project [5], we replicated the studies of [6]. Because our results coincided with hers, we decided to introduce new variables into the simulator, like kind and distribution of food. Also, we carried out a naturalistic observation with a group of seven *Cercocebus torquatus lunulatus* (or *Cercocebus atys lunulatus,* classified among the primates of the old world, in the family of the Cercopitecidae), and compared the empirical data with data provided by the simulator. Observational data was obtained by video recording a group of white crowned mangabeys at the Zoo of Barcelona for over 30 hours. First, we used *The Observer XT* (Noldus Information Technology, Wageningen, The Netherlands) to code the spatial location of every individual at every moment; second, a model of the zoo facility where this group lives was created with *Autocad,* in order to calculate the exact distances among the positions coded with *The Observer*; third, we compared these interdistances with those provided by the simulator at every time step. As a result, of comparing data from the simulations and the naturalistic observation, we obtained a valuable feedback that permitted us to develop a better simulator and to better understand the strategies used by primates to form hierarchical structures.

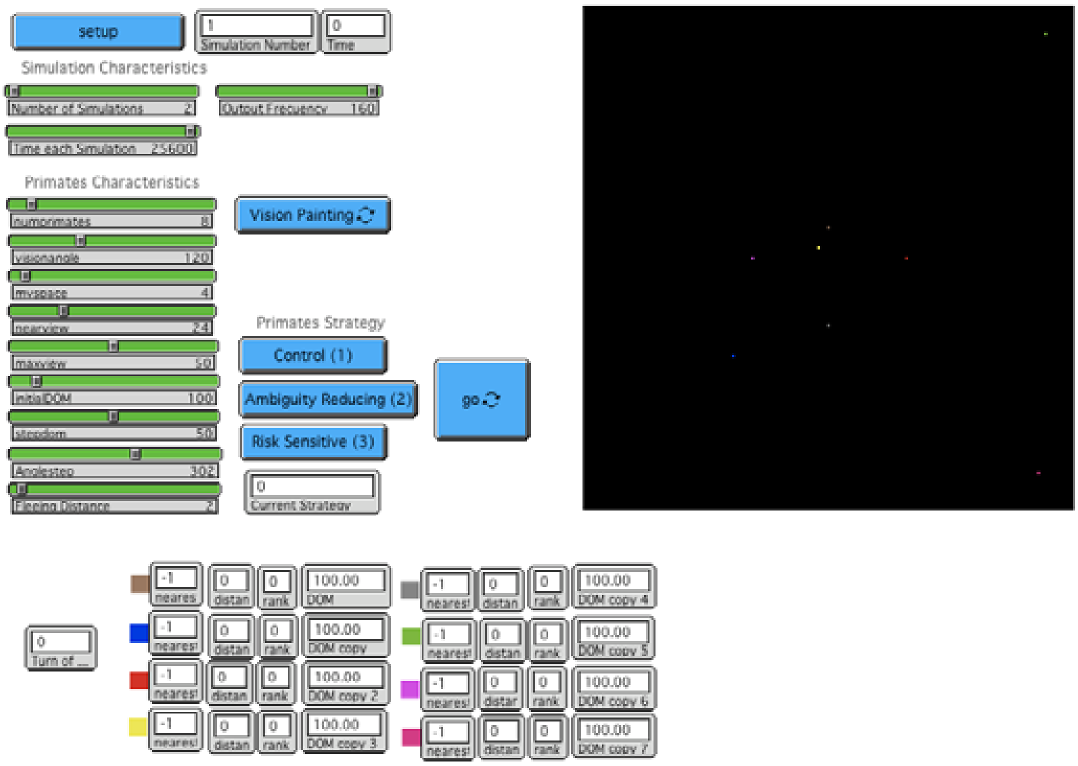


Figure 1. *Screen capture of DOMLogo primate interaction simulator.*

References

1. Koenig, A. (2001). Competition for resources and its behavioral consequences among female primates. *International Journal of Primatology* **23 (4)**, 759-783.
2. Barta, Z. & Giraldeau, L. (1998). The effect of dominance hierarchy on the use of alternative foraging tactics: a phenotype-limited producing-scrounging game. *Behavioral Ecology and Sociobiology* **42**, 217-223.
3. Hemelrijk, C. K. (1998). Risk sensitive and ambiguity reducing dominance interactions in a virtual laboratory. In R. Pfeifer, B. Blunberg, J.-A. Meyer & S. W. Wilson (Eds.). *From Animals to Animats **V*** (pp.255-262). Cambridge, Massachusetts: MIT Press.
4. Colella, V., Klopfer, E., Resnick, M. (2001). *Adventures in modeling : exploring complex, dynamic systems with StarLogo.* New York: Teachers College Press.
5. Cifre, I., Olivé, M., Salvador, F.(2007). Dominància i ocupació de l'espai en grups de primats no-humans: un model basat en conducta adaptativa. *Anuari de Psicologia* **11(1)**, 37-49.
6. Hemelrijk, C. K. (1996). Dominance interactions, spatial dynamics and emergent reciprocity in a virtual world. In P. Maes, M. J. Mataric, J-A Meyer, J. Pollack & S. W. Wilson (Eds.), *From Animals to Animats **IV*** (pp. 545-552). Cambridge, MA: The MIT Press.

Automated detection of compulsive checking behavior in rats

G.G. de Haas[1,2], A. Nijdam[3], T. Westra[1], M.J.H. Kas[2], and H.G.M. Westenberg[1]
[1] *Department of Psychiatry, UMC Utrecht, Rudolf Magnus Institute of Neuroscience, The Netherlands, G.G.deHaas-2@umcutrecht.nl*
[2] *Department of Neuroscience and Pharmacology, Rudolf Magnus Institute of Neuroscience, The Netherlands*
[3] *Department of Animals, Science and Society, Ethology and Animal Welfare,Faculty of Veterinary Medicine, Utrecht University, The Netherlands*

Obsessive compulsive disorder (OCD) is a psychiatric disorder with a prevalence lifetime of 2 to 3%. People with OCD suffer from strong obsessive thoughts and perform rituals (compulsions) to get in control of these obsessions. OCD rituals vary from “washing behavior” (in combination with obsessions related to the fear of being contaminated with germs or dirt) to “checking behavior” or “counting”. Of course checking can be a normal behavior, but OCD patients perform these rituals so excessively that it interferes with their normal daily functioning. Currently, selective serotonin reuptake inhibitors (SSRI's) are the most effective treatment of OCD patients, but still 30-40% does not respond to this medication [1].

Pre-clinical research with valid animal models for OCD is necessary to gain more insight in the pathogenesis of OCD and to discover new therapies. At this point there are no well validated OCD animal models available. Szechtman et al. [2, 3] have preformed research on a candidate animal model in which quinpirole sensitized rats develop compulsive checking behavior. The aim of this study is to investigate the compulsive checking behavior in this model in more detail and evaluate the potential use of Theme [4] for automated detection of behavioral patterns in this animal model.

In this model rats are injected twice a week with the D2/D3 receptor agonist quinpirole 0.5 mg/kg for 5 weeks. After each injection, the animal is placed on a large open field (160X160 cm) with 4 objects and is tracked for 60 min with EthoVision. During the course of treatment the animals will develop compulsive checking behavior. The EthoVision data is fed into the software package Theme (Noldus Information Technology, The Netherlands) which is a tool for pattern detection and analysis in time-based data.

Data analysis in Theme results in a set of hierarchical time patterns, also called T patterns, for individual animals at different time points (sessions). Our data showed that quinpirole treated animals had significantly fewer variations in their behavioral repertoire compared with saline animals (Mixed model; $P < 0.0001$) and that the patterns of their behavioral sequences are less complex. The mean occurrences per pattern showed a significant treatment effect, in which quinpirole treated animals showed a 2-fold higher mean occurrence per pattern compared to control (throughout all sessions), indicating that quinpirole animals repeat their patterns more often (Mixed model; $P < 0.0001$). Most used level and length of patterns are higher in the saline than quinpirole treated animals (Glimmix Poisson; $P=0.0007$, $P=0.0027$), indicating that quinpirole treated animals have shorter and less complex patterns. Related to these results is the mean maximum time of a pattern which is significantly longer in saline treated animals (Mixed model; $P < 0.0001$), which can be explained by the longer length and higher complexity of these patterns. We can conclude that T pattern analysis is a useful tool in the development of an animal model for compulsive behavior. It not only adds a valuable measurement, but also provides a new dimension in the behavioral data that translates to OCD.

References

1. Pallanti, S., Hollander, E., Bienstock, C., Koran, L., Leckman, J., Marazziti, D., Pato, M., Stein, D., Zohar, J., International Treatment Refractory OCD Consortium (2002). Treatment non-response in OCD: methodological issues and operational definitions. *Int J Neuropsychopharmacol.* **5**, 181-91.
2. Szechtman, H., Sulis, W., Eilam, D. (1998). Quinpirole induced compulsive checking behavior in rats: A potential animal model of Obsessive Compulsive Disorder (OCD). *Behavioral Neuroscience* **112**, 1475-1485.
3. Dvorkin, A., Perreault, M.L., Szechtman, H. (2006). Development and temporal organization of compulsive checking induced by repeated injections of the dopamine agonist quinpirole in an animal model of obsessive-compulsive disorder. *Behavioural Brain Research* **169**, 303-311.
4. Magnusson, M.S. (2000). Discovering hidden time patterns in behavior: T-patterns and their detection. *Behavior Research Methods, Instruments & Computers* **32**, 93-110.

Automatic segmentation of mouse behavior using hidden Markov model

D. Vetrov[1], D. Kropotov[2], A. Konushin[1], E. Lomakina-Rumyantseva[1], I. Zarayskaya[3], and K. Anokhin[3]
[1]*Graphics and multimedia lab, Moscow State University, Moscow, Russia, {vetrovd, ktosh}@graphics.cs.msu.ru, lr2kate@gmail.com*
[2]*Dorodnicyn Computing Center of the Russian Academy of Sciences, Moscow, Russia, dkropotov@yandex.ru*
[3]*Department of Systemogenesis, P.K. Anokhin Institute of Normal Physiology, Moscow, Russia, k_anokhin@yahoo.com*

Abstract

A common approach to analysis of mouse behavior recorder by video tracking systems employs manual segmentation and labeling of mouse activity into behavioral acts. Developed automatic methods allow segmentation only to lingering and progression segments, suffer from poor precision and require parameter tuning. We propose a novel approach based on hidden Markov model for simultaneous segmentation and labeling of mouse trajectory into behavior acts. The method uses manually labeled video sequences for training. The developed approach has shown promising results when applied for segmentation of mouse behavior in a novel environment.

Keywords

Video tracking, elementary behavioral acts, activity modeling, hidden Markov models (HMMs).

Introduction

The need for high throughput behavioral screening procedures for rodents has led to introduction of automated home cages that provide various stimuli and equipped with video tracking systems [1]. This, however, results in a drastic increase in complexity regarding observation and analysis. Detailed recordings of position alone yield already an enormous set of data. Advanced analysis techniques, e.g. behavior patterns extraction and evaluation, requires segmentation of tracking data into behavioral acts. Trained human observer is required for this task. Several automatic segmentation methods have been proposed, e.g. [2], but they are restricted to distinguishing lingering and progression segments mostly. Additionally, results of manual segmentation significantly differ between experts. This limits test results replicability between labs and even between different researches in one lab.

However, in field of visual surveillance for human activities a number of approaches have been proposed for event recognition [3]. Similar methods can be applied for distinguishing behavioral acts. We propose a novel method for behavioral act recognition based on hidden Markov models [4].

Proposed method

Our method of trajectory segmentation is based on first order hidden Markov models (HMM) [4]. This approach is an example of sequential probabilistic model and it is widely used in signal segmentation. The system (mouse) being modeled is assumed to be a process with unknown state that generates an output in the form of features, calculated from corresponding mouse trajectory. Each type of behavioral act is considered as a state of the process. The parameters of HMM are estimated from a training set of output sequences with known states. We use manually segmented trajectories with hand-labeled behavior acts as training set. 13 types of behavioral acts were identified by experts. Due to the fact that currently our tracking system can produce only x- and y-coordinates of mouse center of mass trajectory, we combined acts into 4 metastates: «Cleaning» (grooming, climbing, digging, stretching), «Stopping» (quiet, no state, head up), «Active» (running, jumping), and «Searching» (sniffing, turning around, turning head, turning body). These states were formed basing on 4 most probable behavioral acts (mentioned at first position in the description of each metastate) and attaching the remaining ones to the most similar metastate.

We use several features, including speed, acceleration, angle between directions of speed in the current and previous time moment, angle acceleration. The likelihood function for each metastate was computed in the following way: all features were transformed to main axes (projections to eigenvectors of sample covariance matrix) in order to decorrelate them; then we estimated the density of each main axis using the mixture of 5 univariate Gaussians; finally we multiplied probability densities of each main axis to get joint density estimate. The histogram of the first main axis and the corresponding 5-gaussian approximation for metastate «Cleaning» is shown in figure 1.

The segmentation of new trajectory is performed by calculating the most probable sequence of metastates according to features which are computed for each time sample.

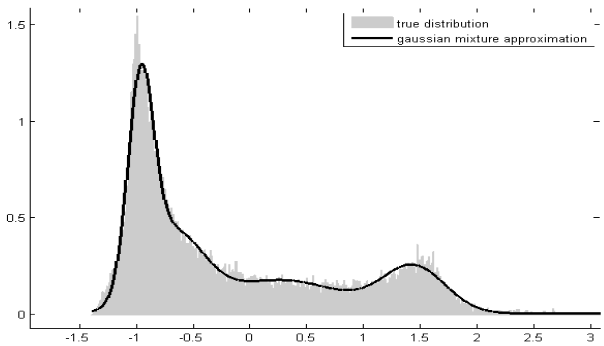


Figure 1. *Approximation of training sample distribution for state Cleaning and the first main axis by mixture of 5 univariate Gaussians. True distribution is shown by grey histogram and Gaussian mixture approximation is shown by black curve.*

Experiments and future work

We have tested proposed method on 13 video sequences of bank vole exploration behavior in open field arena, 325 minutes total. 150 minutes were used as training set, and the rest as testing set. The results of automatic segmentation were compared with the manual ones, see table 1 and figure 2. These results allow us to make the following conclusions: we may classify relatively well metastates associated with fast mouse movements («Active») and with mouse stops («Stopping»). The errors in this classification happen due to the differences in detecting the borders of metastates. Two other metastates («Cleaning» and «Searching») are less distinguishable when we consider only the location of mouse center. It should be noted that different experts segment trajectories in such a way that they differ in average in 15-20% of time points. The last remark allows us to claim that our results are promising. We plan to use additional features from

Table 1. *Confusion matrix for mouse behavior states recognition.*

\From class: Classified as \	«Cleaning»	«Stopping»	«Active»	«Searching»
«Cleaning»	5683	7120	22	3899
«Stopping»	97	102859	95	1883
«Active»	0	21	7590	1796
«Searching»	850	7281	656	10382

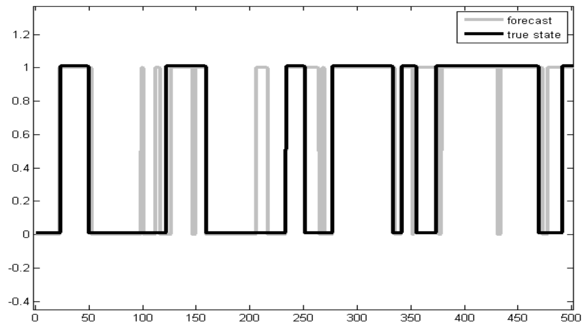


Figure 2. *Example of forecasting mouse trajectory states. Black curve shows right answers, grey curve corresponds to forecasted states values. For simplicity purposes only Stopping (y=1) vs all other metastates (y=0) is shown*

mouse nose and tail base points tracking to improve the accuracy and working with behavior acts directly. Also, higher order HMMs can be used to reduce the small labeling errors.

References

1. Spruijt B.M., DeVisser L.(2006) Advanced behavioral screening:automated home cage ethology. *Drug Discovery Today: Technologies* Vol. **3**, No. 2 2006, pp.231-237
2. Cherepov A.B., Mukhina T.V., Anokhin K.V.(2005) Automatic segmentation of mouse behavior during video tracking in home cages. *5th Int. Conf. on Methods and Techniques in Behavioral Research "Measuring Behavior 2005"*, Wageningen, The Netherlands.
3. J. Aggarwal J., Cai Q.(1999), Human motion analysis: A review. *Comput.Vis. Image Understand.*, vol. 73, no. 3, pp. 428–440.
4. Elliot R.J., Aggoun L., Moore J.B. (1995) Hidden Markov Models: Estimation and Control. *Springer*.

Control of hive environment by honeybee (*Apis mellifera*) in Japan

Mizue Ohashi[1], Hidetoshi Ikeno[1], Toshifumi Kimura[1], Tadaaki Akamatsu[1], Ryuichi Okada[2], and Etsuro Ito[2]
[1]*School of Human Science and Environment, University of Hyogo, Hyogo, Japan, ohashi@shse.u-hyogo.ac.jp*
[2] *Kagawa School of Pharmaceutical Sciences, Tokushima Bunri University, Kagawa, Japan, okd@ncp8.es.hokudai.ac.jp, eito@kph.bunri-u.ac.jp*

Introduction

It is known that honeybee, a kind of social insects, could control hive environment in order to survive in drastic environmental changes in field condition. Colonies of honeybee can keep the temperature within fairly narrow limits over a wide range of outside temperatures (Simpson, 1961). In order to know the thermoregulation mechanisms in honeybee hive, intensive studies have been conducted to correlate the temperature variability with changes in metabolic activity and honeybees' behaviors (e.g. Kronenberg and Heller, 1982). Existence of control mechanisms of humidity and carbon dioxide (CO_2) concentration are also suggested (Seeley, 1974; Human et al. 2006), but little is known about whether they are controlled independently from temperature because controls of temperature normally accompany with changes in respiratory emission of CO_2 and evaporation of hive water and nectar. It is also unclear how the colonial regulation of hive environment varies with diurnal and seasonal changes in ambient environment. Therefore, continuous measurements of multiple factors of hive and ambient environments are likely important in order to reveal the mechanism of environmental regulation of social insects, which differentiate them from other insects. In this study, we aimed to measure temporal changes in temperature, humidity, and CO_2 concentration simultaneously within and out of a honeybee hive. Changes of the hive weight were also measured in order to estimate how much energy consumed for the colony homeostasis.

Material and methods

Measurements have started in December 2007 using a hive of European honeybee (*Apis mellifera*). During the winter (December-February), the hive was placed in a large container (2.6W x 1.4D x 2.0H meters), where light environment was controlled in a fixed cycle (12 hours irradiation every day), but moved into field condition thereafter. We measured hive temperature and humidity using a bandgap and polymer sensor (Seirision; SHT-11) every minute and CO_2 concentration using nondispersion infrared radiation analyzer (Vaisala; GMT220, GMD20) every 10 seconds by 25 January and minute thereafter. These measurements were also conducted in an empty hive and the ambient in order to determine the impact of honeybee on the colonial environment. Hive weight was measured using a weight meter (A and D: FG-60KAL) hourly.

Results and discussion

The weight of hive decreased continuously around 30 g per day during the winter. Temperature within the hive increased in daytime and decreased in nighttime, which was constantly higher than outside. The diurnal pattern had a peak occurring later than that of the ambient, suggesting a time lag of temperature fluctuation between the two locations. The temperature of empty hive had similar tendency to the ambient, but the peak also delayed. This result suggests that the increase of active hive temperature caused by the presence of honeybee, supporting the former studies about the colonial thermoregulation of honeybees (e.g. Simpson, 1961). The humidity within the hive was higher than that of the ambient and fluctuated little even when the ambient humidity changed considerably. Since the magnitude and fluctuation patterns of the active hive were almost same with those of the empty hive, the humidity variation might be caused by the hive material and structure, not by the activity of honeybee. The hive CO_2 concentration fluctuated corresponding to the hive temperature even when atmospheric CO_2 concentration was in stable. We also observed that it sometimes increased drastically in midday, which might be caused by a circadian rhythm of honeybee (Kronenberg and Heller, 1982). In this study, we found that honeybee hive has a different microclimate from those of ambient, which partly caused by the honeybee activity. We also observed that humidity and CO_2 concentration could get effects from their own fluctuation, suggesting another controlling mechanisms for these factors that were different from those for the thermoregulation.

References

1. Human H, Nicolson SW, Dietemann, V (2006) Do honeybees, Apis mellifera scutellata, regulate humidity in their nest? *Naturwissenschaften* **93**: 397-401.
2. Kronenberg F, Heller CH (1982) Colonial thermoregulation in honey bees (Apis mellifera) *J. Comp Physiol. B* **148**: 65-76
3. Seeley TD (1974) Atmospheric carbon dioxide regulation in honey-bee (Apis mellifera) colonies. *J. Insect Physiol*. **20**: 2301-2305.
4. Simpson J (1961) Nest Climate Regulation in Honey Bee Colonie: Honey bees control their domestic environment by methods based on their habit of clustering together. *Science* **133**: 1327 – 1333.

Favourite swimming patterns in the Morris water maze: allothetic and idiothetic navigation in small rodents

Michala Zelenkova[1] and Frantisek Sedlacek[1,2]
[1]University of South Bohemia, Faculty of Sciences, Ceské Budejovice, Czech Republic, zelenm02@prf.jcu.cz
[2]Institute of Systems Biology and Ecology AS CR, Ceske Budejovice, Czech Republic, sedlacek@usbe.cas.cz

Introduction

Animals apply two navigation systems to reach their goal by the shortest way in familiar habitats: allothetic and idiothetic. The allothetic navigation system is based on the determination of direction and distance to the goal according to the relationships between used landmarks animals can find in their habitats [1]. On the other hand, in the idiothetic navigation system animals rely on the vestibular system, proprioreceptors and the muscle spindles as the sources of information [2]. Both of these systems can be investigated in the Morris water maze (MWM) [3]. The most of studies about navigation abilities in MWM are based on the experiments with laboratory rats and mice. Hence we chose three wild rodent species, *Microtus arvalis*, *Acomys dimidiatus*, *Mus musculus,* and one laboratory mouse (outbred CD-1 strain) that underwent classical MWM training. We wanted to make better sense of rodent searching strategies which could be one of reasons they are able to find their goal successfully, so we evaluated observed repetitious observed swimming patterns for the first, third and fifth day of testing.

Methods

Our MWM consists of a pool (95 cm in diameter, 50 cm in height) surrounded with a non-transparent and non-translucent tent. Every animal had to swim eight trials per day from four start positions to find a submerged escape platform which we placed in one fixed position in the pool. Their path length was recorded. We compared the allothetic navigation of the rodents, when three accentuated and contrast landmarks were placed inside the tent, and their idiothetic navigation, when no landmarks were present.

Evaluated strategies often occurred in all species or they were special for one species, more or less. During thigmotaxis, animals were swimming near the wall because lots of animals are afraid to swim far from the wall. When animals left the wall and made an arch to return, we denominated this swimming as arch near the wall. Animals explored more parts of the pool but they didn't swim to the centre. We found some individuals that swam across the whole area of the pool. Thus we also recorded swimming across. During floating, animals were lying still on the water surface. Sometimes animals were searching for the platform in one part of the pool and their path resembled big circles. On the other hand, small loops, whereby animals were searching through the whole pool, were typical for semicircular swimming. Some animals didn't regard the platform as a safe place and they abandoned it immediately, so the recording of their path didn't stop. When animals headed for the centre and they missed the platform but soon they returned to it, we described this pattern as return. We also noted the frequency of direct swimming: In the end we recorded the frequency of the following swimming patterns: thigmotaxis, arch near the wall, swimming across, floating, big circles, semicircular swimming, abandoning the platform, return, direct.

Results

All species were able to find the platform more accurately in the presence of visual landmarks than without them. All species preferred semicircular swimming for idiothetic navigation, whereas we didn't find any swimming patterns most preferred for allothetic navigation. The abilities of *Acomys dimidiatus* to find the platform were worse than of all the other species. This result is consistent with our presumption that the natural environment affects the navigation abilities of species. Moreover, *Acomys dimidiatus* also preferred swimming patterns such as thigmotaxis, arch near the wall and floating more than other species. These patterns might be interpreted as a bad searching strategy. We suggest that in addition to the relation between the natural environment of the species and their ability to orientate with or without landmarks, the preferred strategy for searching platform can affect the performance of small rodents in MWM.

References

1. Benhamou, S., Poucet, B. (1996). A comparative analysis of spatial memory processes. *Behavioural Processes* **35**, 113-126.
2. Etienne, A.S., Jeffery, K.J. (2004). Path integration in mammals. *Hippocampus* **14**, 180-192.
3. Moghaddam, M., Bures J. (1996). Contribution of egocentric spatial memory to place navigation of rats in the Morris water maze. *Neurobiology of Learning and Memory* **78**, 121-129.

Measuring the neuronal activity of hypothalamus and behavior of rats on the diffuse low gravity conditions

Y. Kumei[1], J L. Zeredo[2], S. Seki[1], M. Matsuura[3], M. Kimoto[4], D. Kageyama[5], Y. Fusejima[6], T. Ikeda[2], and K. Toda[2]
[1]*Graduate School of Tokyo Medical and Dental University,* [2] *Graduate School of Nagasaki University,* [3]*Japanese Foundation for Cancer Research,* [4]*Japan Women's University,* [5]*Diamond Air Service, Co.,* [6]*Japan Space Forum Foundation, Japan*
kumei.bch@tmd.ac.jp

The purpose of the present study is to examine the rat behavior on diffuse low gravity conditions. We have succeeded for the first time in generating diffuse low gravity conditions from 0.4G through 0.01G by using the unique method of parabolic flights that were operated by Diamond Air Service, Co at the Nagoya airport in Japan. The final goal is to assess the potential risk and/or benefit for the animal and human activity on the low gravity conditions such as those encountered on the surface of the Moon and Mars. In parallel with the observation of rat behaviors, the neuronal activity in the hypothalamic arcuate nucleus (ARC) was measured by telemetry through chronically implanted microelectrodes in freely moving rats during the parabolic flights. All the rats were maintained in the facilities on a 12-hr light/dark cycling schedule with ad libitum access to food and water throughout the duration of the experiment. The treatment of the rats was in accordance with the guidelines of Tokyo Medical and Dental University, Nagasaki University, and the animal use committee of JAXA Japan Space Agency. The insertion site of the electrodes was identified histologically following rat sacrifice after all the experimental operations were completed post-flight. As the airplane produces low gravity conditions by a parabolic vertical flight path, the aircraft will exert less G force as relative to the aircraft. Shortly after the aircraft was heading upward at a certain angle, the pilot immediately started the rotation into the parabolic trajectory. While the aircraft flies on this locus, it loses altitude significantly to generate the targeted low gravity level for several seconds. The aircraft finally pulls into an upward posture to get back to the normal static flight. Four rats were used aboard the aircraft daily, and a total of 12 rats were used during four continuous days. Rats were exposed to a certain level of low gravity repeatedly for three times in each day. Rat behavior was recorded by digital video cameras throughout the static and parabolic flights at varied gravity conditions. Cowering was observed in rats that were exposed to 0.2G and 0.15G. Locomotor activity, grooming, freezing, depression, and rearing up were observed at varied low gravity levels from 0.15G through 0.05G. Tumbling, a typical sign that is observed in weightlessness, was observed at 0.01G. These stress-related behaviors were observed less frequently by the repeated exposure to the same level of low gravity. The rat “adaptation” to low gravity environments was assessed quantitatively by comparing the neuronal activity of the same rat in response to the 1^{st} time exposure versus to the 3^{rd} time exposure of the same gravity condition (paired t-test). Low gravity was the stressful condition for rats, of which activity was quantified by the neuronal activity at the neuroendocrine control center, hypothalamus. In conclusion, we have succeeded for the first time in assessing rat behaviors in diffuse low gravity conditions. Our success in the unique flight experiments would provide beneficial impacts on the future life sciences in differential gravity environments. The achievements in this study are useful to predict the human activity on the Moon and Mars in the future manned space flight missions.

This study was supported by Japan Space Forum Foundation to Y.Kumei (FY2006-2007) and by a Grant-in-Aid for Scientific Research from the Japan Society for the Promotion of Science to Y.Kumei (No. 18659569, FY2006-2007).

A comparison of different methods for quantifying stereotypic behaviour patterns in captive European starlings (*Sturnus vulgaris*)

Lucy Asher[1], Ben Brilot[2], Gesa Feenders[2], and Melissa Bateson[2]
[1]*Royal Veterinary College, Hatfield, UK, lasher@rvc.ac.uk*
[2]*Centre for Behaviour and Evolution, Newcastle University, UK, Melissa.Bateson@ncl.ac.uk*

Stereotypies, defined as behaviour patterns that are repetitive, unvarying and with no apparent function, are common in laboratory animals. For example, many caged birds show route-tracing stereotypies in which an individual repeatedly visits the same locations, and those familiar with the behaviour of caged starlings will have observed the "somersaulting" stereotypy in which birds perform backward aerial flips repeatedly (Figure 1). Such behaviour patterns have been described in a wide range of captive mammalian and avian species including farm, zoo, companion and laboratory animals. They are most common in barren and or spatially restricting cages and are often interpreted as a sign of poor welfare.

Figure 1. *Composite photo showing a somersaulting starling.*

Standard techniques for quantifying stereotypies involve counting of individual incidences of a route being traced or a distinctive behaviour pattern such as a somersault. However, defining these complex sequences of behaviour objectively can be difficult, and is likely to fail to capture stages in the development of stereotypies while sequences of behaviour are still flexible[1]. For these reasons we are developing measures of stereotypy that are both more objective and have the potential to capture early developmental stages of stereotypy.

We will present a comparison of three different techniques for quantifying stereotypies from videotapes of the behaviour of caged European starlings (*Sturnus vulgaris*):

1. Manual counting of complete somersaults (the most easily recognized stereotypy observed in caged starlings).
2. Use of Markov chains to quantify the sequential dependency in a sequence of behaviour. This technique is a modification of one described by Haccou[2], and uses the probabilities of one event following another (P(Z/Y) (first order) or another two (P(Z/XY) (second order) to describe and quantify the sequential dependency in a sequence of behaviour. The more repetitive a sequence of events, the more events Y-Z will be preceded by X. We have applied this technique both to sequences of individual behaviour patterns (e.g. hop, preen, probe substrate etc) and to sequences of spatial locations in the cage occupied by the bird (e.g. floor, perch 1, front wall etc). The latter technique proved much simpler and less time consuming to apply than the former, and also has the benefits of being easier to totally automate. We have already demonstrated that this technique picks up significant differences between starlings housed in cages of different sizes and shapes for as little as one week (see Figure 2).
3. A pattern recognition algorithm based on T-pattern analysis implemented in the commercially available software package, Theme. T-pattern analysis is capable of finding patterns in sequences of behaviour that would be invisible to Markov chains analysis, but has the disadvantage of being harder to interpret. This technique has been used to demonstrate significant differences in the pattern of performance on a gambling task in stereotypic human patients suffering from schizophrenia compared with healthy controls[3]. There is also one abstract suggesting that Theme differentiates mice treated with different doses of the dopamine transporter inhibitor GBR-12909, a drug known to induce locomotor stereotypies[4]. To date, the only application of Theme in a welfare context is to compare behavioural complexity in broilers on different feeding regimens[5].

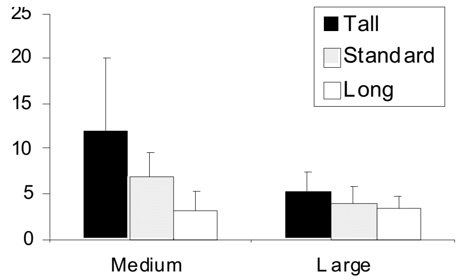


Figure 2. *Pilot data showing the effects of cage size and shape on the degree of stereotypy (sequential dependency derived from Markov chains analysis) in starling behavior.*

Preliminary analyses show that those animals displaying full somersaults are those with high sequential dependency scores[6], suggesting that our Markov chains approach may be effectively detecting birds with higher levels of stereotypy.

We will use the three methods described above to investigate:

1. The initial development of stereotypies in recently-caught starlings transferred to small cages for the first time.
2. The incidence of stereotypies in starlings housed in enriched versus barren cages.

References

1. Meehan, C.L., Garner, J.P., and Mench, J.A., (2004). Environmental enrichment and development of cage stereotypy in Orange-winged Amazon Parrots (*Amazona amazonica*). *Developmenal Psychobiology* **44**, 209-218.
2. Haccou, P., and Meelis, E., (1992). *Statistical analysis of behavioural data; an approach based on time structured models*. Oxford Univeristy Press, Oxford.

Proceedings of Measuring Behavior 2008 (Maastricht, The Netherlands, August 26-29, 2008)
Eds. A.J. Spink, M.R. Ballintijn, N.D. Bogers, F. Grieco, L.W.S. Loijens, L.P.J.J. Noldus, G. Smit, and P.H. Zimmerman

3. Lyon, M., and Kemp, A.S., (2004). Increased temporal patterns in choice responding and altered cognitive processes in schizophrenia and mania. *Psychopharmacology* **172**, 211-219.
4. Bonasera, S.J., Schenk, K.A., and Tecott, L.H., (2006) T pattern sequential analysis to quantify "route tracing" locomotor stereotypic behavior, *Society for Neuroscience Abstracts*, Georgia World Congress Center.
5. Hocking, P.M., Rutherford, K.M.D., and Picard, M., (2007). Comparison of time-based frequencies, fractal analysis and T-patterns for assessing behavioural changes in broiler breeders fed on two diets at two levels of feed restriction: a case study. *Applied Animal Behaviour Science* **104**, 37-48.
6. Asher, L., (2007) *The welfare of captive European starlings (Sturnus vulgaris)*. PhD Thesis, Institute of Neuroscience, Newcastle University, Newcastle upon Tyne, UK.

Combined use of the PhenoTyper® and accelerometer for activity and sleep pattern categorisation

Andrea Plano, Gernot Riedel, and Bettina Platt
Institute of Medical Sciences, University of Aberdeen, UK, bms343@abdn.ac.uk

Rationale

Sleep pattern analysis requires the recording of both EEG and EMG in order to categorise different sleep stages; however the EMG recording has technical problems due to anatomical and physiological limits of the animals. Therefore, a more tailored alternative would be beneficial for an accurate categorisation of sleep stages. The PhenoTyper® system provides useful information of gross locomotor activity in a free moving animal in its home cage [1]. However, the software (EthoVision 3.0) only detects robust changes of the body point centre, and it does not record fine movements (such as small head accelerations) during the stationary phase. The knowledge of possible subtle movements is fundamental for sleep stages categorisation as it is a hallmark to distinguish the different sleep phases and during active and quiet awake. The use of an accelerometer would provide such information [2], however limitations would occur if for instance, an animal moves with a constant velocity, moreover it does not provide information on the animal spatial location. The combined use of the PhenoTyper and an accelerometer device would provide a more robust reading of the activity of the animal, which would allow a more reliable interpretation of EEG recordings during circadian rhythms.

Methods

Twelve month old female C57BL6/j mice were anesthetized with 3% isoflurane in medical grade oxygen. After full induction of the anaesthesia gold electrodes were placed in the prefrontal cortex and left and right hippocampi. Reference ground screws were placed in the parietal and occipital regions. Electrodes in form of gold skull-screws, soldered to pins, were lowered 1-1.5 mm below the skull to touch the dura mater. Finally, the electrode assembly of seven pins was anchored and fixed to the skull. Once mice recovered from surgery, (at least 7 days post surgery) behavioural experiments took place.

Mice were connected with a wireless EEG recorder data logger with a built-in accelerometer and their activity was synchronically recorded for 1 hour during the light phase by both the PhenoTyper® (EthoVision 3.0, Noldus Information Technology, Wageningen, The Netherlands) and the built-in accelerometer. The device would also provide EEG recordings from selected brain areas, but this was not included in the present experiment. Animals were visually observed during the test to combine the software and visually recorded data. "Active awake" and "quiet awake" status was defined by the experimenter as active awake status with high activity and quiet awake with little or no movements.

Results

During both active and quiet awake assessed by visual observation and the use of the PhenoTyper®, a correlation between the locomotor activity recorded by EthoVision and the acceleration recorded by the data logger was found (see Figure 1 and 2). Correlations was obtained during both active ($P<0.001$, $R^2=0.01912$, slope= 0.5584 ± 0.09) and quiet awake phases ($P<0.001$, $R^2=0.1584$, slope= 15.66 ± 0.9799). However, during the active phase, the accelerometer reached the maximal measurable acceleration as it becomes saturated whereas EthoVision continually provides linear information about the activity. During the quiet wakefulness the animal stays still but sometimes grooms, moves the head or sleeps so in this case the accelerometer is able to provide useful information in order to discriminate between different stages of quiet activity.

Conclusions

In the active phase (see figure 1), the accelerometer saturates rapidly whereas EthoVision provides complete information about the animal activity. In the quiet phase (see figure 2) there is a wide distribution of the parameters, therefore for a small range of movements the accelerometer gives a more detailed reading of small changes in activity typical of pre-sleep phases. This is fundamental for the automatic scoring analysis of sleep stages. In conclusion, the PhenoTyper® (EthoVision 3.0) allowed a gross analysis of the locomotor activity, whereas the accelerometer provided an index of fine accelerations of the head. Together the two devices offer an enriched reading of the animal's activity allowing a more robust interpretation of the automatic recognition of the behavioural parameters and possibly sleep stages.

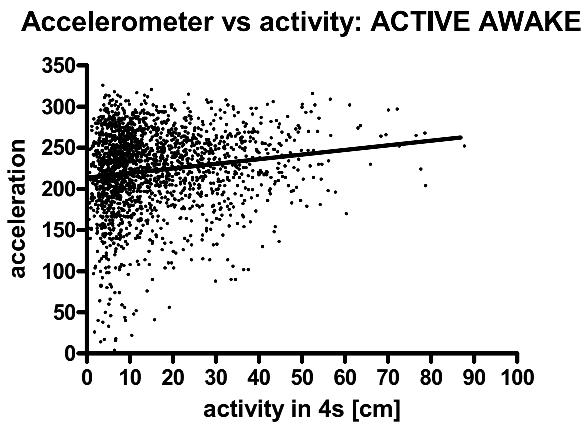


***Figure 1**. Linear interpolation between active activity and acceleration*

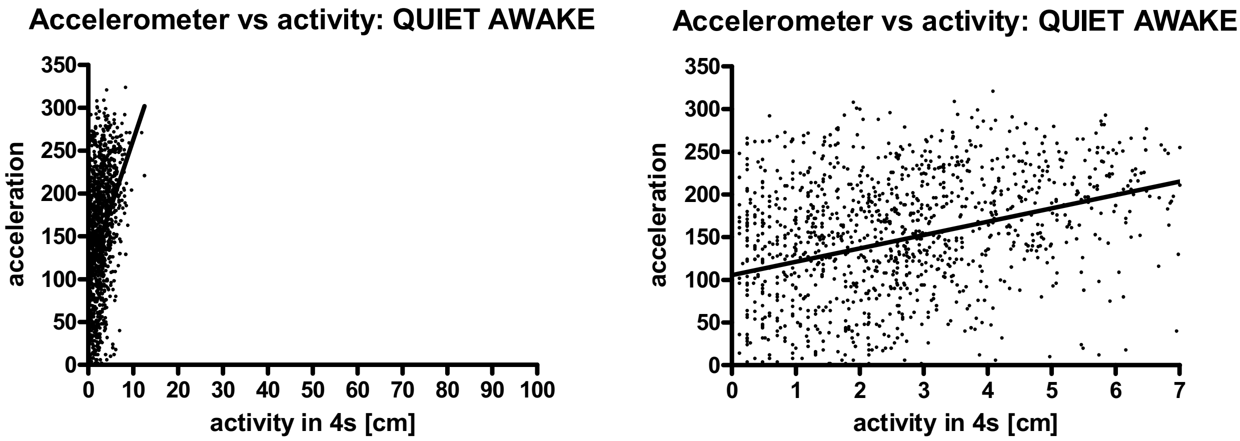


Figure 2. *Linear interpolation in different scales between quiet activity and acceleration*

References

1. Visser, de, L.; Bos, van den, R.; Kuurman, W.; Kas, M.; Spruijt, B. (2006). Novel approach to the behavioural characteristics of inbred mice: automated home cage observations. *Genes, Brain and Behavior,* **5***, 458-466*
2. Sunderam S., Chernyy N., Peixoto N., Masona J.P., Weinstein S.L., Schiff S.J., Gluckmana B.J.. (2007). Improved sleep–wake and behavior discrimination using MEMS accelerometers. *Journal of Neuroscience Methods* **163,** 373–383.

A new automated system to study discriminative learning for auditory cues

V. Lioudyno, A. Kiryk, R.K. Filipkowski, and L. Kaczmarek
Nencki Institute of Experimental Biology, Warsaw, Poland, vlioudyno@mail.ru

The ability to discriminate spectral and temporal characteristics of auditory stimuli is an important factor for animals in coping with the environment. However, studying of these processes in laboratory rodents is associated with technical problems. Until now the main behavioral paradigm to study auditory discrimination in laboratory rodents was differential fear conditioning which is strongly aversive and could not be performed in a home cage. To deal with this challenge we have designed a new behavioral paradigm to study ability of mice to discriminate between auditory stimuli using the automated Novel Tune Cage system (New-Behavior; http://www.newbehavior.com). In this paradigm, a group of mice lives in a regular cage connected by a tunnel to a sound-proof chamber (see Figure 1A). Inside the chamber, there is an antenna reading the individual transponder code assigned to each mouse. The shape of the access tunnel restricts the entrance to the chamber for a single mouse only. The water bottles are available only for mice in the chamber, and the animals have to perform a nose-poke in order to open small motorized doors; the nose-pokes are registered by photo sensors.

After a few days of adaptation to the regular cage, the learning procedure with auditory stimuli was started. Tone bursts with 8 kHz frequency ("standard sound") were presented when the mouse was recognized by the sensor during 4 days. Next 4 days two tones were presented randomly. The 75 % of visits were accompanied by the same standard tone and 25 % with another one, "deviant". Deviant was presented either as continuous tone of 11.4 kHz (first experiment, n = 11 mice) or as tone bursts with 11.4 kHz (second experiment, n = 9 mice). For the both experiments each nose-poke performed during the presentation of the deviant tone was punished by an air-puff. At the time of standard sound any nose-poke was opening door and providing access to water. So, depending on the sound presented different module either opening the doors and allowing to drink or activating the punishing air-puff was turned on (see Figure 1B, C).

No changes in general activity of mice were observed after introduction of the standard sound in both experiments. The first "deviant" visit of each mouse was always followed by a nose-poke reaction and hence it was always punished by an air-puff. But when standard and deviant sounds differed in frequency as well as in temporal structure (first experiment) each mouse had only a few deviant visits finished with a nose-poke at the end of the first day of the discriminative procedure. Also, average duration of the deviant visits was significantly lower comparing to the duration of "standard" visits. These results have shown that an unknown sound when paired with aversive stimulus, such as the air-puff, becomes significant for the mice. The animals start to recognize the novel sound as unsafe, differentiate it from the familiar sound and, as a result, develop appropriate behavioral strategy to avoid the punishment. Moreover, it appeared that the temporal structure of auditory signal may be very important for mice, as in second experiment (when two sounds had the same temporal structure and were different only in frequency) only 5 from 9 mice in a group were able to discriminate between the auditory cues. The rest being unable to escape punishment stopped to perform nose-poke reaction in time of deviant as well in time of standard sound soon after discriminative learning procedure was started. So, analysis of general activity shown phenomena of generalization for those mice who had no success in discriminative task.

We conclude that Novel Tune Cage is an effective behavioral system to study auditory discriminative learning in mice. The system allows also revealing the individual differences in learning capacity and may be useful to study the impact of particular genes on recognition of auditory cues.

The study was supported by 6FP Novel Tune project.

A B C

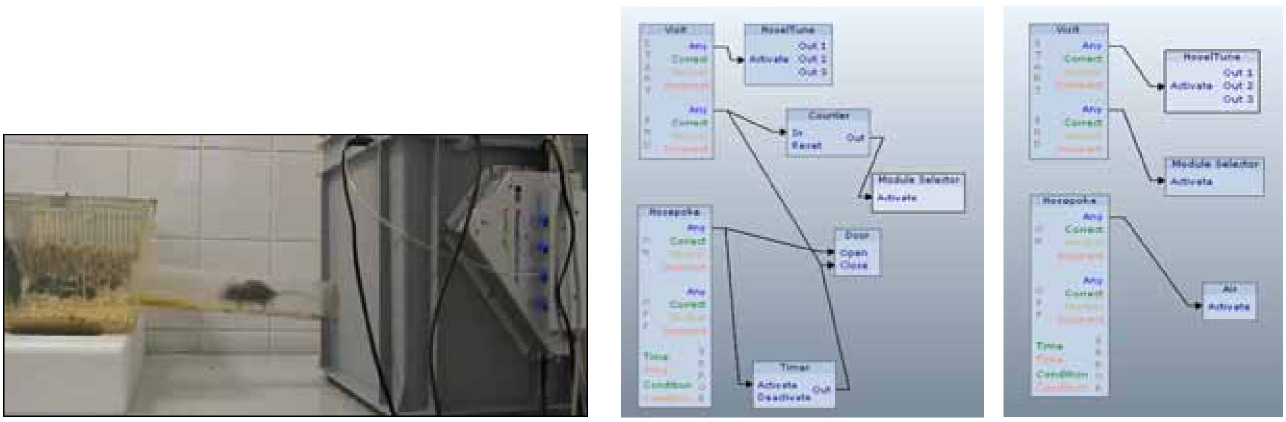


Figure 1. *The general view of Novel Tune Cage (A); modules for "standard" (B) and "deviant" (C) visits during the discrimination learning procedure.*

Motor assessment in transgenic Tau mice

Valeria Melis, Gernot Riedel, and Charlie Harrington
Institute of Medical Sciences, University of Aberdeen, Aberdeen, UK, men238@abdn.ac.uk

Tau is a microtubule-associated protein functionally known to bind microtubules and to be critical in the process of neuronal outgrowth and axonal integrity [1-4]. Intracellular accumulation of filamentous tau inclusions are neuropathological hallmarks of neurodegenerative disease known as tauopathies which are associated with the progressive loss of cognitive, behavioural and motor functions [5-7]. The discovery that mutations in the tau gene can cause a familial form of fronto-temporal dementia with parkinsonism linked to chromosome 17, (FTDP-17) has established the central role of tau dysfunction in neurodegenerative disease [8,9]. Several transgenic mouse lines over-expressing mutant human tau in neurons and glial cells have now been expressed [10].

The aim of this study was to provide a behavioural characterization of one such transgenic mouse line (line 66) as a potential model for investigating the role of tau in neurodegeneration. These transgenic mice, bred on a NMRI background strain, exhibit a progressive and age-related deterioration in motor performance and learning skills, features common in the tauopathies. One phenotypes of line 66 mice is a palsy-like shaking or wobbling of the body which emerges at about 7-9 months of age. Line 66 mice were allocated to two different groups depending on whether they visibly displayed this phenotype ("wobblers") or not ("non wobblers"). Using a battery of sensorimotor tasks, we investigated the differences in sensorimotor skills, gait and motor learning for line 66 groups compared with wild-type control mice. The following tests were used: the Rotarod test to assess motor learning and motor coordination; the balance beam test to examine sensorimotor coordination and balance of mice and an automated Catwalk analysis to provide a large number of gait parameters (both static and dynamic). In the Rotarod task, mice were placed on an accelerating rod and the latency to fall from the rod recorded for a maximum period of 300 sec. In the balance beam test, mice had to climb six beams differing in both thickness (28-, 11- or 5-mm) and shape (square or round) in a maximum time of 30 sec. The latency to reach the upper end of a beam, at inclination of 30°, was the dependent measure. In the Catwalk mice had to traverse from one end to the other of a glass plate. With this computer-assisted method of locomotor analysis, it is possible to quantify several gait parameters, including the duration of different phases of the step cycle and the pressure applied during locomotion.

Using the Catwalk system, we have quantified a large number of gait parameters for line 66 and wild-type mice. Marked groups differences were noted on the static parameters; the hindlimb base-of-support was wider in wild-type mice compared to line 66 groups. A feasible explanation for the observed differences might be that line 66 mice tend to hold their hind legs in a crossed position close to their bodies. Also, as shown in relative paw placement results, line 66 have a propensity to place their hind paws in front the previously placed ipsilateral forepaws. Inter-limb coordination is a key characteristic of locomotion and therefore the parameters relating to coordination are of particular interest. All strains used the two alternate patterns (i.e. Aa and Ab) and the two cruciate patterns (i.e. Ca and Cb) with a preference for the Ab pattern. In particular, the higher proportion of wobblers using the Ab step sequence pattern compared with wild type mice might be as accounted for by an adjustment in walking balance. The shaking of the body, in line 66 wobblers, makes it extremely unstable for the mouse to walk with a simultaneous swing of one body side and so the use of left front and right hind paws followed by right front and left hind paws (Ab) represents the most stable way for these mice to walk. Rotarod and balance beam tests revealed that line 66 mice developed significant motor impairments compared with non transgenic age-matched control mice. It was possible to also detect motor and sensory motor differences within the same genotype. In both tests, wobblers showed a worse performance relative to the non wobbler group indicating a higher deficit in coordination and balance skills.

Line 66 mice, therefore, offer relevant information as a simplified model for human tauopathies and, together with other model systems, it may assist in our understanding of the role of tau in neurodegeneration and in the development of therapeutic agents.

References

1. Brandt, R., Lee G. (1993).The balance between tau protein's microtubule growth and nucleation activities: implications for the formation of axonal microtubules. *J. Neurochem*. **61**, 997–1005.
2. Brandt, R., Lee G. (1993). Functional organization of microtubule-associated protein tau. Identification of regions which affect microtubule growth, nucleation, and bundle formation in vitro. *J. Biol. Chem*. **268**, 3414–3419.
3. Cleveland, D.W., Hwo, S.Y., Kirschner, M.W. (1977). Physical and chemical properties of purified tau factor and the role of tau in microtubule assembly, *J. Mol. Biol.* **116,** 227– 247.
4. Weingarten, M.D., Lockwood, A.H., Hwo, S.Y., Kirschner, M.W. (1975). A protein factor essential for microtubule assembly, *Proc. Natl. Acad. Sci. USA* **72**, 1858–1862.
5. Spillantini et al. (1997). Familial multiple system tauopathy with presenile dementia: a disease with abundant neuronal and glial tau filaments, Proc Natl Acad Sci U S A. **94,** 4113-8.
6. Perry, R.J., Hodges J.R. (1999). Attention and executive deficits in Alzheimer's disease: a critical review. *Brain* **122**, 383-404.
7. Swartz, J.R., Miller, B.L., Lesser, I.M., et al. (1997). Behavioral phenomenology in Alzheimer's disease, frontotemporal dementia, and latelife depression: a retrospective analysis. *J Geriatr Psychiatry Neurol.***10**, 67-74.
8. Poorkaj, P., Bird, T.D., Wijsman, E., Nemens, E., Garruto, R.M., Anderson, L. et al. (1998).Tau is a candidate gene for chromosome 17 frontotemporal dementia. *Ann Neurol*. **43,** 815–25.
9. Spillantini, M.G., Crowther, R.A., Kamphorst, W., Heutink, P., van Swieten, J.C. (1998) Tau pathology in two Dutch families with mutations in the microtubule-binding region of tau, *Am. J. Pathol.* **153**, 1359–1363.
10. Lewis, J., McGowan E., Rockwood, J., Melrose, H., Nacharaju, P., Van Slegtenhorst, M., Gwinn-Hardy, K., Paul Murphy, M., at al. (2000). Neurofibrillary tangles, amyotrophy and progressive motor disturbance in mice expressing mutant (P301L) tau protein, *Nat. Genet*. **25,** 402–405.

T-patterns analysis differentiates between the open field behavioral strategies in two species of wild rodents

I.Yu. Zarayskaya and K.V. Anokhin

P.K. Anokhin Institute of Normal Physiology of RAMS, Moscow, Russia, irzar2003@mail.ru

Introduction

Detailed studies of rodent exploratory behavior demonstrated that they establish space structure of routes and places in a novel environment, including a home base and excursions from it and back to it [1]. It has been found that when a rat or a mouse is placed in a new environment (e.g. open field apparatus) it produces two main forms of activity: forward progression and lingering. Lingering episodes are involved in investigation of a particular location, while forward progression brings the animal from one location to the next [2]. Quantifying them separately allows to differentiate between exploratory behavior of different species and strains of rodents [3]. However, attempts to reveal various strategies of rodent exploratory behavior based only on algorithmic segmentation and subsequent classification of locomotor behavior into different motion modes run into difficulty of lacking information about successions of discrete behavioral acts from which the “forward progression” and “lingering” episodes consist of.

Used approach

The aim of the present study was to reveal a natural structure of rodent exploratory behavior in the open field test as consisting of hierarchically organized patterns of behavioral acts. Our approach to identification of discrete acts was based on the development of track segmentation algorithms [4] that permit to divide a behavioral continuum into the single behavioral episodes and to detect connection between them. In the present study we performed comparative analysis of exploratory behavior in two species of wild rodents: bank voles (*Clethrionomys glareolus*) and pygmy wood mice (*Sylvaemus uralensis)*. We studied their behavior in an open field apparatus (d=120 cm) in two 15-min test sessions separated by 24 hours. Video tracking was performed by “Easy Track” software developed in our lab [5]. The subsequent segmentation of behavioral continuum into single behavioral acts was performed by custom developed “Segment Analyzer” software [4]. In total we isolated 14 different behavioral categories: “running”, “quiet”, “sniffing”, “stretching”, “head turn”, “body turn”, “turn around”, “head up”, “rear support”, “rear”, “jump”, “climbing”, “digging” and “grooming”. Obtained successions of discrete behavioral acts were then analyzed by the «Theme» software to reveal the structure of T-patterns [6].

Experimental data and further work

During both open field tests pygmy wood mice (*Sylvaemus uralensis*) displayed a broader spectrum of behavioral act categories as compared to the bank voles (*Clethrionomys glareolus*). However, analysis of behavior patterns revealed clear similarities between sequences of acts in the two species under both (first and second tests) conditions. Primarily it concerned “body turn-sniffing” and “turn around-sniffing” behavioral sequences. These patterns were involved in investigation of particular locations and therefore were labeled by us as “lingering” patterns, analogously to the category from the motion mode analysis experiments [3]. Patterns “body turn – quiet”, “head turn - quiet”, “turn around - quiet” reflected the resting behavior. These patterns were connected to home bases in both species. “Forward progression” episodes contained “lingering” acts in all cases. Such patterns as “running - rearing with support”, “body turn -running” and “running-sniffing” were found in each studied animal.

Major differences between the two species were found in the “lingering” strategies, which were mostly connected to the investigation of particular locations. Patterns «body turn - digging», «body turn - rearing», «rearing with support - climbing», «rearing with support - digging», «sniffing - digging» were found only in pygmy wood mice, while «head turn – body turn», «head turn - sniffing» and «head turn – turn around» were present in bank voles only.

We conclude that proposed approach to analysis of open field behavior structure highlights peculiarities of exploratory behavior organization in the studied rodent species. In future we intend to test the hypothesis that the found T-patterns of open field behavior can be fixed actions programs of rodent exploratory behavior.

References

1. Eilam D, Golani I. (1989). Home base behavior of rats (*Ruttus norvegicus*) exploring a novel environment. *Behavioral Brain Reseach, I,* **34**:199-211.
2. Kafkafi, N., Mayo, C. L., Drai, D., Golani, D. & Elmer, G. I. (2001). Natural segmentation of the locomotor behavior of drug-induced rats in a photobeam cage. *Journal of Neuroscience Methods,* **109**, 111-121.
3. Drai, D. & Golani, I. (2001). SEE: a tool for the visualization and analysis of rodent exploratory behavior. *Neuroscience and Biobehavioral Reviews*, **25**, 409-426.
4. Cherepov A.B., Mukhina T.V., Anokhin K.V. (2005). Automatic segmentation of mouse behavior during video tracking in home cages. *Proceedings of Measuring Behavior 2005, 5th International Conference on Methods and Techniques in Behavioral Research*. Ed. by L. Noldus, F. Grieco, L. Loijens and P. Zimmerman. Wageningen: Noldus Information Technology. P. 403.
5. Anokhin K. and Mukhina T. (2005). Development of a neurocognitive interface for an animal video tracking system.; *Proceedings of Measuring Behavior 2005, 5th International Conference on Methods and Techniques in Behavioral Research*, Ed. by L. Noldus, F. Grieco, L. Loijens and P. Zimmerman. Wageningen: Noldus Information Technology.
6. Magnusson M.S. (1996). Hidden real-times patterns in intra- and interindividual behavior: Description and detection. *Eur. J. of Psychological Assessment* **12**, 112-123.

Development of automatic analysis and recognition of mouse behavior by segmentation and t-pattern method using video tracking

A. Cherepov and K. Anokhin
Department of Systemogenesis, P.K. Anokhin Institute of Normal Physiology, Moscow, Russia, ipmagus@mail.ru

Introduction

Automated recognition of human and animal behavior is a rapidly developing trend in behavioral researches. This recognition is of great practical importance (e.g. total video monitoring in airports and railway stations, screening researches of new drugs, etc.) and is a key element in fundamental studies of brain functions and behavior.

Since 2002 we developed a new approach to quantitative segmentation of animal behavior by automatic dissection of behavioral continuum into meaningful behavioral units [1] and developed a novel video tracking system [2]. Our approach described elsewhere A.B. Cherepov, et al. [3] is based on the functional systems theory of Peter Anokhin (FST) [4]. Our first results were presented at previous Measuring Behavior meeting [3]. The best segmentation was obtained by application of running median to find stops followed by the analysis of acceleration projection for obtaining segments in order to find points representing additional breaks and, hence, additional acts. Analysis of the video records of the segmented behavior demonstrated that the breakpoints found by this method showed the best match to the segmentation performed by an expert observer (about 85% coincidences).

The present research was conducted to develop algorithms for automatic recognition (classification) of behavioral units in locomotor behavior of mice.

Methods and results

In order to reveal automatically different classes of behavioral acts we used approach proposed by Drai et al. 2001 [5]. After primary dissection of the locomotor behavior into behavioral segments and stops between them, the distribution of the maximum speed was analyzed using Gaussian Mixture Model (GMM). This method models the distribution as a sum of several normal distributions. In our experiment, the distribution was typically bimodal, two types of behavioral segments are thus revealed (Type I and II). Type I behavioral segments (low velocity) mainly corresponded to rearings, stretchings, sniffings, groomings, etc. Information obtained by tracking of the center of mass only is insufficient to recognize each type of such low-velocity behavioral acts (Type I). Below we discuss only recognition of Type II behavioral acts. Normally, Type II acts include only running acts, but Gaussian distributions in GMM usually considerably overlap and some rearings, stretchings, and groomings can be erroneously included into Type II group. To solve this problem we applied an additional criterion named “bias threshold”. Bias is the shortest distance between the start and end points of the act (Figure 1). If this shortest distance calculated for the act erroneously identified as “Type II” was below the bias threshold (3-4 cm in our tests) it is automatically included into the correct group (Type I).

In order to test the proposed algorithms we analyzed behavior of C57Bl/6 mice either under amphetamine treatment (acute administration i.p. 5 m/kg) in home cages or placed in a novel environment (new “home” cages of the same size with fresh bedding and transparent walls, “novelty group”). Mouse activity was recorded using our video tracking software. Segmentation of behavior was performed with modified version of SegmentAnalyser software containing comprehensive options for breakpoint detection, classification and visualization [4]. These experiments have shown the best match to segmentation into behavioral units performed by an expert observer (about 85% coincidences) for both behavior groups of mice (d-amphetamine-treated and placed in novel environment).

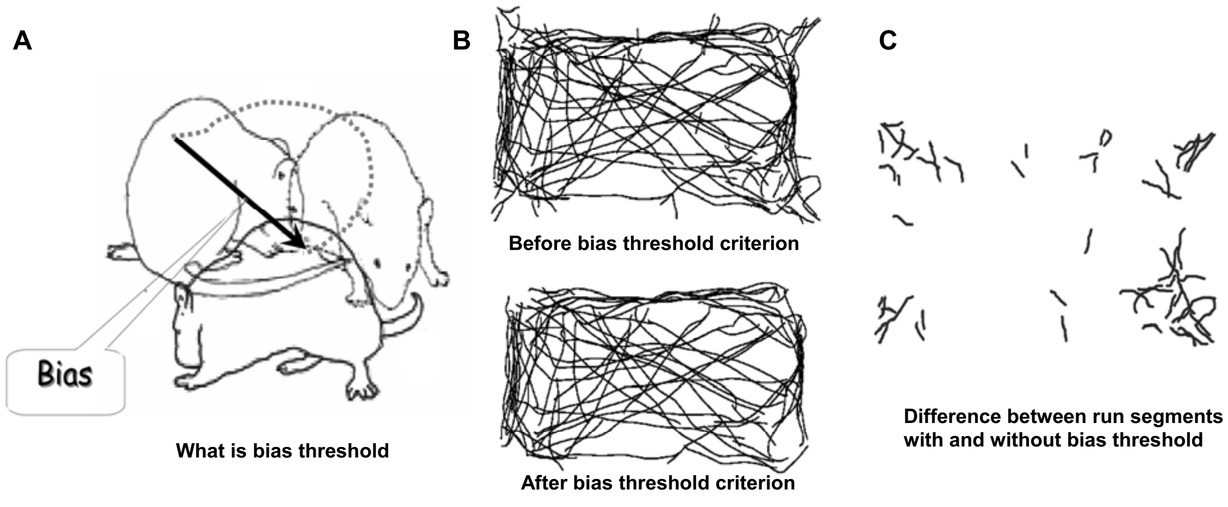


Figure 1. *Bias threshold algorithm.*

Conclusions

Bias threshold criterion allows us to correct automatic discrimination between runnings and low-speed acts, which results in 85% coincidence in automatic and expert observer's classification of behavior units.

This new algorithm is appropriate for two different mouse behavior types, and probably can be applied to analyze other types of behavior.

Ethograms obtained by expert observers were analyzed using t-pattern analysis (Theme 5.0) [6]. Behavior repertoire consisted of ten behavior units: climbing, digging, grooming, quiet, rear, rear with support, run, sniff, stretch, turn around. Two most meaningful t-patterns, "sniff-run" and "turn around-run", were detected in the behavior of both mouse groups. These t-patterns constitute about 90% of all observed patterns. Practically all runs performed by mice were components of these t-patterns. Thus, mouse behavior in these behavioral tasks can be precisely characterized by detection (manual or automatic) of all runs. And our method does allow precise automatic detection of runs!

References

1. Mukhina, T.V.; Lukashov, A.O.; Anokhin, K.V.; Bachurin, S.O. (2002). Analysis of animal path trajectories in terms of individual behavioral acts. *4th Int. Conf. on Methods and Techniques in Behavioral Research "Measuring Behavior 2002"*, Amsterdam, The Netherlands. 27-30 August, 2002.
2. Anokhin, K. and Mukhina, T. (Moscow, Russia). Development of neurocognitive interface for animal video tracking system. *5th Int. Conf. on Methods and Techniques in Behavioral Research "Measuring Behaviour 2005"*, Wageningen, The Netherlands 2005
3. A.B. Cherepov , T.V. Mukhina, K.V. Anokhin (Moscow, Russia). Automatic segmentation of mouse behavior during video tracking in home cages. *5th Int. Conf. on Methods and Techniques in Behavioral Research "Measuring Behaviour 2005"*, Wageningen, The Netherlands 2005
4. Anokhin P.K. (1968). Nodular mechanisms of functional system as self-regulatory apparatus. *Progress in Brain Research*. **22,** 230-251.
5. Drai, D.; Golani, I. (2001). SEE: a tool for the visualization and analysis of rodent exploratory behavior. *Neuroscience and Biobehavioral Reviews*. **25**, 409-426.
6. Magnusson MS (2000) Discovering hidden time patterns in behavior: t-patterns and their detection. *Behav Res Meth Instrum Comput* **32**, 93–110

An automated home cage observation system as a model of feeding behaviour in mice

Lianne Robinson, Susan McKillop-Smith, and Gernot Riedel
School of Medical Science, College of Life Science and Medicine, University of Aberdeen, Aberdeen, Scotland, UK, l.robinson@abdn.ac.uk

Introduction

Obesity is characterised by an increase in consumption of food containing high levels of sugar and saturated fats combined with a reduction in physical activity. The increased prevalence of obesity and childhood obesity in western countries has highlighted the need for anorectic agents which are effective in weight management.

The Cannabinoid (CB1) antagonist SR141716A (Rimonabant/Acomplia) has previously been shown to suppress food intake [1, 2]. However, in addition to drug induced effects on food intake and body weight, drug treatment can also affect activity levels [3, 4]. An increase in activity may lead to hyperactivity which could explain a reduction in body weight which is independent of any reduction in food consumption. Therefore, a more detailed assessment of drug-related effects on activity in conjunction with drug-induced hypophagic properties is required.

Aim

The aim of this study was to develop a novel method which would combine these two approaches and allow us to assess both the feeding behaviour and locomotor activity of free-feeding mice. The cannabinoid antagonist AM251 was used to validate this method.

Methods

Male and female C57BL6 mice (25-32g) were used in this study. The automated home cage video-based observation system 'phenotyper' was used to allow long-term continuous monitoring of behavioural activity of the mice. During testing the mice were maintained on a 12 hr light/dark cycle (lights off 7pm) with free access to food (food hopper) and water (water bottle). Animals were given 3-4 days of habituation in the phenotyper prior to drug administration in order to attain a baseline level of performance and were matched for body weight before being assigned to drug groups of AM251 (10 mg/kg) or Tween 80, with all drugs injected intraperitoneally 1-2 hours prior to the start of the dark phase of the circadian cycle. A single acute treatment or repeated dosing regime was employed with the drug administered daily for 4 consecutive days. The behaviour of the animals was monitored using the computer based tracking software Ethovision (version 3.0, Noldus, Netherlands). The parameters recorded included total locomotor activity, time spent in the area in front of the food hopper and time spent in the area close to the water bottle. These parameters were analysed in hourly intervals during both the light and dark phases of the experimental days. In addition to these parameters the body weight of each animal, weight of food hopper (food intake) and weight of water bottle (water intake) were recorded on a daily basis immediately prior to drug injection and also post drug treatment (between the hours of 10.00 – 11.00 am of the light cycle).

Results

Acute treatment with AM251 induced a suppression of food intake and a reduction in body weight compared to vehicle treated animals. AM251 treated animals also spent less in the areas of the arena associated with food and water. In addition a reduction in locomotor activity was also observed although this was only apparent for the first few hours of the dark phase. Repeated dosing with AM251 also produced a significant suppression in body weight and food intake relative to controls, with animals spending less time in the food zone on all nights following treatment. Similar to acute treatment a decrease in overall locomotor activity was also evident.

Conclusions

These findings suggest that the home cage observation system 'phenotyper' is a sensitive and effective method to assess the hypophagic effects of possible anti-obesity drugs as it allows the long-term continuous monitoring of both food intake and behavioural activity.

References

1. Colombo, G, Agabio, R, Diaz, G, Lobina, C, Reali, R, Gessa, GL (1998). Appetite suppression and weight loss after the cannabinoid antagonist SR 141716. *Life Science* **63**: PL113-117
2. Vickers, SP, Webster, LJ, Wyatt, A, Dourish, CT, Kennett, GA (2003). Preferential effects of the cannabinoid CB1 receptor antagonist, SR141716, on food intake and body weight gain of obese (fa/fa) compared to lean Zucker rats. *Psychopharmacology* **167**: 103-111
3. Compton, DR, Aceto, MD, Lowe, J, Martin, BR (1996). In vivo characterization of a specific cannabinoid receptor antagonist (SR141716A): Inhibition of Delta (9)- tetrahydrocannabinol-induced responses and apparent agonist activity. *Journal of Pharmacology and Experimental Therapeutics* **277 (2)**: 586-594
4. Bass, CE, Griffin, G, Grier, M, Mahadevan, A, Razdan, RK, Martin, BR (2002). SR141716A induced stimulation of locomotor activity: a structure-activity relationship study. *Pharmacology Biochemistry and Behavior* **74**: 31-40

Inhalation of neroli essential oil and its anxiolytic effects in animals

Ying-Ju Chen[1,3], Ying Shih[2], Tsong-Min Chang[2], Ming-Fu Wang[3], Sen-Sen Lan[1], and Fu-Chou Cheng[1]
[1]*Stem Cell Medical Research Center, Department of Medical Research, Taichung Veterans General Hospital, 407 Taichung, Taiwan fucheng@vghtc.gov.tw, lansensen@gmail.com*
[2]*Department of Applied Cosmetology, Hung Kuang University, 433 Taichung, Taiwan, krystle@sunrise.hk.edu.tw; ctm@sunrise.hk.edu.tw*
[3]*Departmentof Food and Nutrition, Providence University, 433 Taichung, Taiwan, g9122501@gmail.com; mfwang@pu.edu.tw*

Current studies have shown that vapors of essential oils such as neroli and lavender could help reduce anxiety levels [1]. Inhalation of these essential oils may trigger the limbic system of the brain giving the resulting beneficial effects. However, further investigation of these oils is needed to provide evidence that they relieve anxiety [2]. This study was carried out with gerbils subjected to aromatherapy inhaling neroli. Forced swimming task and locomotor activity were measured to evaluate levels of anxiety [3,4]. Neroli essential oil used in this study was supplied by Australian Botanical Products Ply Ltd. (Hallam Victoria, Australia). 100 µL of neroli was vaporized in a nebulizing diffuser and was inhaled by individual gerbils for 30 min in the behavior box. The neroli-inhaled group received 100 µL neroli essential oil overnight before taking a forced swimming task. During the exercise, a few drops of neroli were mixed in the water for the gerbils to inhale. Locomotor activity was measured by means of a multi-box ActiMot detection system (TSE, Bad Homburg, Germany). Two gerbils were simultaneously placed in two open field boxes for control and neroli-inhaled groups. The boxes were covered with cloth and the room was kept dark and silent to prevent any interference from the outside environment. One minute of habitation was given prior to the 30-minute measurement of activity. Each gerbil was placed on a polystyrene board floating in a Plexiglas cylinder that was 15 cm in diameter, 35 cm in depth, and was filled with warm water (*ca.* 35°C) to a height of 18 cm. Total duration of swimming was measured during the experiment. Gerbils were judged to be immobile when they floated passively with the head above water.

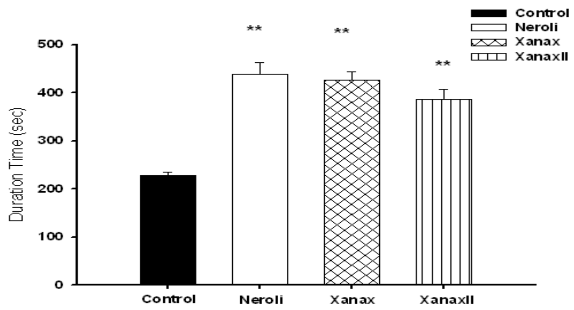


Figure 1. *Comparison of swimming duration time in the control and neroli groups (**p<0.01).*

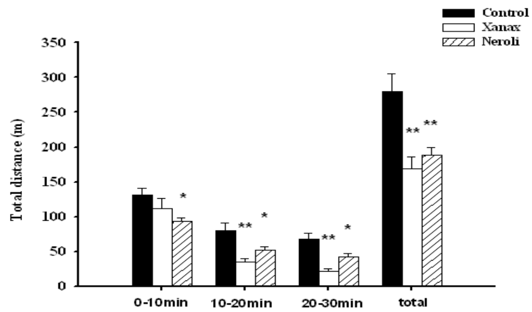


Figure 2. *Comparison of total distance in the control and treatment groups (*p<0.05 and **p<0.01).*

Comparison was made between duration of forced swimming tasks and total distance between the control and neroli-inhaled groups. In addition, the treatment of Xanax®, an anxiolytic drug, was used as a positive control. The average duration times for swimming were 228±7.0, 439±23, 386±21, and 427±18 seconds in the control, neroli-inhaled, and two Xanax-treated groups, respectively (see Figure 1). These treatments were significantly increased by 65-91% of basal in all groups ($p<0.01$) when compared with the control. The total distance traveled during 30 min were 280.3±25.4 m, 189.0±10.7 m and 168.6±17.6 m in the control, neroli-inhaled, and Xanax group, respectively (see Figure 2).

The forced swimming task was designed to screen anxiolytics and antidepressants. Immobility observed during the task was described as a lower state of mood in animals. This passive behavior caused by stress is seen as an adaptive response in an inescapable situation. In the present animal study, swimming times were increased significantly (65% and 82% of the basal in Xanax and XanaxII groups, respectively). Xanax® is a short-acting drug of the benzodiazepine class used to treat anxiety disorders, panic attacks and nervous tension, different types of phobias, depression, insomnia and others. It works by slowing down the nervous system and controlling chemicals in the brain. From these results we concluded that Xanax® could act as a positive control in gerbils subjected to a forced swimming task. It was also observed that swimming times were increased by 91% of the basal using neroli when compared with the control. These effects of neroli are comparable to both Xanax and control groups in our results. Previous studies have shown that inhalation of essential oil caused a decrease of immobility in mice using essential oils. The present study showed that inhalation of neroli vapor may induce anti-anxiety effects. Side effects of Xanax® are most likely an extension of the pharmacological activity of alprazolam including sleepiness, memory problems and blurred vision in patients. In addition, Xanax® has been known to be habit-forming. Withdrawal symptoms may occur if treatment is discontinued abruptly. On the contrary, the essential oil has reduced anxiety in mice without generating side effects in previous reports.

According to results of the forced swimming task, significant increases in duration time were observed in the neroli-inhaled group and a reduction of immobility in gerbils was shown to confirm the anxiolytic effects of neroli. To identify the possible neural or hormonal mechanisms in this essential oil, a specifically designed study can be planned. Both gerbils treated with neroli and Xanax® displayed anxiolytic effects by the major decreases in total distances comparing with the control. Aromatherapy for anti- anxiety has been developed based on conventional trials of essential oils, including neroli, by scientists. In general, they found aromatherapy "comforting" or "relaxing and enjoyable". However, the elucidation of anxiolytic effects and the quantitative information for these essential oils have been limited. To date there have been few studies addressing the effects of aromatherapy or massage on anxiety. The present data provides an additional evidence for neroli in the potential treatment of anxiety. Improvements of effectiveness were shown in the inhalation of neroli and Xanax® in both behavior tests. However, the mechanisms of anxiolytic effect responsesfor neroli and Xanax® might be different. This study provides an evidence based information on aromatherapy in the treatment of anxiety.

References

1. Lehrner, J., Marwinski, G., Lehr, S., Johren, P., Deecke, L. (2005) Ambient odors of orange and lavender reduce anxiety and improve mood in a dental office. *Physiol Behav.,* **86**, 92-95.
2. Bradley, B.F., Starkey, N.J., Brown, S.L., Lea, R.W. (2007) Anxiolytic effects of Lavandula angustifolia odour on the Mongolian gerbil elevated plus maze. *J. Ethnopharmacol.,* **111**, 517-525.
3. Petit-Demouliere, B., Chenu, F., Bourin, M. (2005) Forced swimming test in mice: a review of antidepressant activity. *Psychopharmacology (Berl.),* **177**, 245-255.
4. Cannizzaro, C., Plescia, F., Gagliano, M., Cannizzaro, G., Mantia, G., La Barbera, M., Provenzano, G., Cannizzaro, E. (2008) Perinatal exposure to 5-methoxytryptamine, behavioural-stress reactivity and functional response of 5-HT1A receptors in the adolescent rat. *Behav Brain Res.*, **186**, 98-106.

The Effects of Locus coeruleus inactivation on I.V self administration of morphine and morphine withdrawal symptoms in rats

M.Amini

Dep of Physiology , Tehran Azad University, Iran, Mohamad1353@yahoo.com

Biochemical, behavioral, and electrophysiologic studies indicate that activation of the noradrenergic cells in the Locus coeruleus (LC) play an important role in the symptoms of opiate withdrawal . There for in this study the effects of LC inactivation on self-administration of Morphine and on morphine withdrawal syndrome in rats has been investigated.

Male rats (250-300gr) were anaesthetized and implanted with silastic catheters inserted in to the right jugular vein. After 5 days animals were fitted and the external end of the catheter was connected with a syringe-driven pump,then were placed in the self-administration apparatus that had two lever (active,passive) for 2 h every day. Active lever switched on the infusion pump for 10 sec,injecting .1ml of saline or saline containing 5 mg/ml of morphine .(training period was 10days). LC was inactivated by (1ul) lidocaein (%2) five min befor training.

Animals were allowed to self administer morphine (1mg/kg per inf.) ten consecutive daily 2-h session. During all morphine self administration session lever pressing behavior was measured . Our results show that LC inactivation 5min before morphine self administration produced a significant decrease in the initiation of morphine self administration during all session. After the last test session morphine withdrawal symptom signs (MWS) precipitated by naloxone were measured . Our results show that most of MWS (but not all) were decreased by L C inactivation in comparison with morphine group.

5-Fluorouracil counters melatonin-induced alterations in locomotor activity

L. Cassim and S. Daya

Neuroscience Research Group, Faculty of Pharmacy, Rhodes University, Grahamstown, South Africa, S.Daya@ru.ac.za

Introduction

5-Fluorouracil (5-FU) is a commonly-used antineoplastic agent associated with a wide range of toxicities, particularly to the gastrointestinal, immune and haematological systems [1]. Clinical studies have shown that the co-administration of the antioxidant and pineal hormone melatonin to cancer patients on 5-FU therapy both significantly enhances the quality of life of these patients and attenuates 5-FU-induced toxicity, leading to an increase in patient survival time [2]. Having previously reported that melatonin counters 5-FU-induced decreases in brain neurotransmitter levels [3], and given the paucity of information available on the effects of 5-FU on locomotor activity, we decided to investigate the effects of these drugs on a variety of locomotor activity parameters. Alterations in locomotor activity are very common in cancer patients [4, 5], who are known to have lowered endogenous levels of melatonin [6]. Such an investigation would offer insight into whether 5-FU potentially exacerbates these locomotor activity changes, contributing to the development of "chemotherapy-related malaise" [7], for example, and whether co-therapy with melatonin might counter this.

Method

Thirty-two male Wistar rats, each weighing between 200g and 300g, were divided into four treatment groups of eight rats per group: (i) control, receiving 0.9% saline; (ii) 5-FU, at a dose of 1 mg/ kg; (iii) melatonin, at a dose of 1 mg/ kg; and (iv) a group receiving both drugs, each at a dose of 1 mg/ kg. The animals were injected three times a day for five days. Locomotor activity was measured prior to drug treatment, for a twelve hour period of constant bright light, from 07h00 to 19h00, using the Noldus Ethovision® video-tracking system. The parameters investigated were: (i) total distance moved; (ii) maximum distance moved; (iii) mean velocity; (iv) total duration of movement; and (v) rearing frequency. Locomotor activity was similarly measured subsequent to the five days of drug treatment. One-way analysis of variance, followed by the Student-Newman-Keuls multiple range test, was used to determine if there were any statistically significant differences in the pre- and post-treatment locomotor activities of rats across the groups.

Results and discussion

Melatonin administration decreases rearing frequency and the total duration of movement, possibly due to the benzodiazepine-like sedative effects of the compound [8, 9]. The ability of melatonin to decrease the frequency of rearing, an indicator of aggressive behaviour [10, 11] suggests that melatonin has anti-aggressive effects. The administration of 5-FU alone significantly increases the locomotor activity of rats exposed to twelve hours of constant bright light, evidenced by increases in the total distance moved, total duration of movement and mean velocity. This excitatory effect suggests that 5-FU may induce anxiety and/ or depression in the clinical setting. This could be a behavioural effect of the alteration in neurotransmitter levels that we have reported before, in particular due to a fall in serotonin, dopamine and norepinephrine levels [3]. The co-administration of 5-FU to rats receiving melatonin treatment abolishes the effects of melatonin on locomotor activity. This implies that whilst melatonin co-therapy is potentially useful in cancer patients, to normalise physiological parameters and enhance patients' quality of life by strengthening impaired circadian rhythms, inducing sleep, and exerting anti-depressant, anti-aggressive and anxiolytic effects, it is necessary to administer this agent on a chronic basis in order to effectively counteract the adverse effects of 5-FU.

Acknowledgements

Financial support from the Andrew Mellon Foundation and National Research Foundation is gratefully acknowledged.

References

1. Bayomi SM, Al-Badr AA (1989). Analytical profile of 5-fluorouracil. *Analytical Profiles of Drug Substances* **18**: 599-639
2. Lissoni P, Barni S, Mandala M, Ardizzoia A, Paolorossi F, Vaghi M, Longarini R, Malugani F, Tancini G (1999a). Decreased toxicity and increased efficacy of cancer chemotherapy using the pineal hormone melatonin in metastatic solid tumour patients with poor clinical status. *European Journal of Cancer* **35**: 1688-1692
3. Cassim L, Maharaj DS, Maharaj H, Daya S (2006). Melatonin counters the 5-fluorouracil-induced decrease in brain serotonin and dopamine levels. *Annals of Neuroscience* **13**: 27-30
4. Buchanan RJ, Barkley J, Wang S, Kim MS (2005). Analyses of nursing home residents with cancer at admission. *Cancer Nursing* **28**: 406-414
5. Onuma E, Tsunenari T, Saito H, Sato K, Yamada-Okabe H, Ogata E (2005). Parathyroid hormone-related protein (PTHrP) as a causative factor of cancer-associated wasting: possible involvement of PTHrP in the repression of locomotor activity in rats bearing human tumor xenografts. *International Journal of Cancer* **116**: 471-478
6. Mormont MC, Waterhouse J (2002). Contribution of the rest-activity circadian rhythm to the quality of life in cancer patients. *Chronobiology International* **19**: 313-323
7. Malik NM, Moore GBT, Smith G, Liu Y-L, Sanger GJ, Andrews PLR (2006). Behavioural and hypothalamic molecular effects of the anti-cancer agent cisplatin in the rat: a model of chemotherapy-related malaise? *Pharmacology, Biochemistry, and Behavior* **83**: 9-20
8. Golombek DA, Cardinali DP (1993). Melatonin accelerates reentrainment after phase advance of the light-dark cycle in Syrian hamsters: antagonism by flumazenil. *Chronobiology International* **10**: 435-441
9. Shaji AV, Kulkarni SK (1998). Central nervous system depressant activities of melatonin in rats and mice. *Indian Journal of Experimental Biology* **36**: 257-263
10. McGregor IS, Hargreaves GA, Apfelbach R, Hunt GE (2004). Neural correlates of cat odor-induced anxiety in rats: region-specific effects of the benzodiazepine midazolam. *Journal of Neuroscience* **24**: 4134-4144
11. Dielenberg RA, McGregor IS (2001) Defensive behavior in rats towards predatory odors: a review. *Neuroscience and Behavioral Reviews* **25**: 597-609

Challenges for Applying Visual Monitoring Techniques on Small Mammals in Arid Zones

M. A. Al-Mutairi

Agriculture and Biodiversity Department, Food Resources and Marine Sciences Division, Kuwait Institute for Scientific Research, Kuwait, State of Kuwait, mmutairy@kisr.edu.kw

The combination of harsh arid environments and small animals' size forms major constraints on visual tracking of rodent's behavior. In the state of Kuwait, the most abundant mammals in the desert are rodents of small size, and the majority of them are nocturnal. Therefore, night vision should be considered when using visual tracking techniques. In addition to rodents, there are approximately 14 species of mammals recorded in Kuwait [1], most of them were extinct or on the edge of extinction due to habitat loss [2]. Therefore, there is an urgent need to develop and apply the latest techniques to monitor and measure different animals' behavioral activities. If these techniques were proven to be useful and applicable in different environmental conditions, we can eliminate or minimize the causes for the rapid decline in the number of desert mammals.

The aim of the study is to test video camera traps and modify it to record nocturnal behavior of lesser Jerboa (*Jaculus jaculus*). The passive camera trap uses the combination of animal movement and body heat to trigger the video camera to record such activity for certain period of time in a design similar, but not exact, to [3]. Most captured animals using theses cameras were of relatively large sizes that could be easily detected. Lesser jerboa is a small rodent of approximately 300g [4] and it is distributed throughout different arid zones in the Middle East. The questions that needed to be answered are: Can camera traps detect animals of such small size as the lesser jerboa with its fast movement? And can these traps recognize its body heat within the hot and dry conditions? Would the cameras be triggered to operate by the movements of small shrubs that dominate the Kuwaiti desert?

In order to answer these questions, two sets of camera traps (TM700V Trail monitor) were placed in two areas with different vegetation community [5] at both the center and north of the State of Kuwait in attempts to capture the nocturnal feeding behavior of lesser jerboa. The trail monitor was first tested in different locations and several configurations to identify the optimal placement for the device to detect small desert rodents. The device was set to operate at night using infrared light and it would trigger a video camera to record any event that might set the device on including the movements of other animals or simply the movement of plants parts by wind. It was also tested in different environmental conditions such as high temperature, strong wind and sandy storms.

After two months of trails, it was proven that trail monitor using passive infrared lights and triggered by the combination of body heat and animal's movement could be used in desert conditions. The results show different animal species caught on video cameras performing their natural activities without human interfere. Animal species such as; mammals, lizards and even birds were recorded feeding on different plant species or simply using them as shelter. Sometimes, harsh weather conditions were an obstacle in obtaining good and clear recording because they change the camera trap configuration or cause false triggering of the camera.

References

1. Delima E., Peacock, J. & Khalil, E. (2002). Bibliographic list of the wildlife fauna of Kuwait.Technical Report. Kuwait Institute for Scientific Research, Kuwait.
2. Al-Sidrawi, F. (1986). Conservation and protection of the wildlife in Kuwait (phase I): review and assessment of information on desert fauna. Final Report. Kuwait Institute for Scientific Research, Kuwait.
3. Scheibe, M., Eichhorn, K., Wiesmayr, M., Schonert, B. & Krone, O. (2008). Long-term automatic video recording as a tool for analyzing the time patterns of utilisation of predefined locations by wild animals. *European Journal of Wildlife Research* **54**:53–59.
4. Clayton, D. (1983). Mammals. Kuwait's Natural History: An Introduction (eds D. Clayton & C. Pilcher, Kuwait Oil Company, Kuwait.
5. Zaman, S., Peacock, J., Al-Mutairi, M., Delima, E., Al-Othman, A., Loughland, R., Tawfiq, H., Siddiqui, K., Al-Dossery, S. & Dashti, J. (2005). Wildlife and vegetation survey (Terrestrial Cluster 3) Addendum 2. Fianl Report, 7885. Kuwait Institute for Scientific Research, Kuwait.

Assessment of learning strategy in young and aged rats in modified holeboard test: Prevention of memory deficits by ladostigil

T. Poltyrev and M. Weinstock

Department of Pharmacology, School of Pharmacy, Hebrew University, Jerusalem, Israel, tatyanapo@ekmd.huji.ac.il

Aim

The aim of the study is to investigate learning strategy in young and old rats, to detect the leading modalities in formation of behaviour in holeboard test [1] (HB) and to assess the potential of ladostigil (TV-3326) - a novel drug with anti-inflammatory activity [2] to prevent aged-related cognitive decline.

Methods

The subjects were male Wistar rats, aged 3.5 months (Young, n=10) and 20 months (Aged, n=14; Aged + ladostigil, n=8). Chronic oral treatment with ladostigil (1 mg/kg) commenced 4 months before training and lasted till the end of testing. *Apparatus:* The holeboard is a box (70 * 70 * 45 cm) with dark brown transparent walls and lid. There are 16 holes (3.5cm diameter) in the black opaque floor. Each hole with a plastic cup (1.5 cm deep) could be illuminated from below. Sunflower seeds (45 mg) served as reinforcement. To minimize odour cues additional seeds (100g) were spread under the floor. The experiments were carried out under dim illumination.

Adaptation sessions began one day after the start of food-restriction. The animals received 85% of their usual ration till the end of the training. The seeds were scattered randomly on the floor and their number was reduced to four at the end of the sessions. Adaptation lasted 5 consecutive days and was considered complete when the rat found 4 seeds during 5 min.

Acquisition of discrimination lasted 9 days with 2 trials per day. Inter-trial interval was 30 min. The rats were trained to collect seeds from a fixed set of four illuminated holes. The trial was terminated when the rat found all the seeds or when 180 sec had elapsed. Visits to the holes were recorded by means of a computer program, which was used to identify nose pokes by light beam crossings. Two measures of training were analyzed:

- Reference memory (RM) = (number of food rewarded visits + revisits to the baited holes) / (number of visits + revisits to all holes)
- Serial numbers of visits to the baited holes (the optimum performance is when the fourth rewarded visit has the serial number of 4)

To assess the contribution of stimuli associated with reward four probe tests were conducted after completion of Acquisition stage. Test 1 aimed to reveal a role of spatial factors. All of 16 holes were illuminated. No food reward was given. The test was terminated when the rat visited all 4 formerly learned holes. Because light in this test has an incentive but not an informative value this probe was called "no food / no light". In Test 2 we investigated the role of (food + light) in directing learned behaviour, apart from its association with fixed location. Four randomly chosen holes were lit and baited with seeds. In Test 3 ("no food/ light") we assessed the role of the learned light signal alone. Four randomly chosen holes were lit. No food reward was given. In Test 4 ("food/ no light") the directing role of odour was investigated. Four randomly chosen holes were baited with seeds. Light was not presented. The last 3 tests were terminated when the rat has visited all 4 holes designated by light and/or odour of food, or when 180 sec had elapsed.

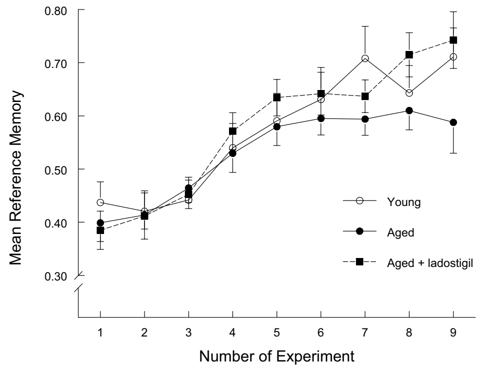


Figure 1. *Reference memory*

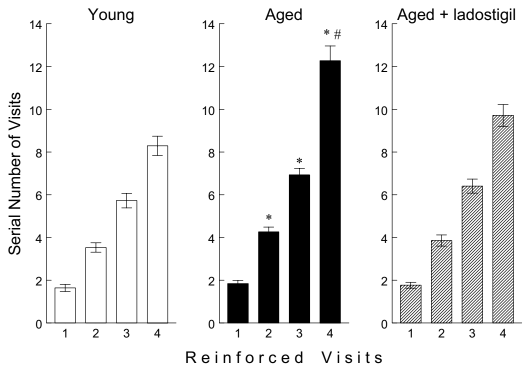


Figure 2. *Serial number of visits to the baited holes *Sig cf Young, p<0.01; # sig cf Aged + ladostigil, p<0.01.*

Results

As a result of food restriction the aged animals lost weight at the end of experiments, but no more than 7% (Young: from (392 ± 9) g to (399 ± 3) g; Aged: from (595 ± 11) g to (572 ± 11) g; Aged + ladostigil: from (635 ± 11) g to (591 ± 7) g. At the stage of Acquisition the rats demonstrated the same rate of initial learning, but in contrast to Young and Aged + ladostigil, Aged rats did not improve their level of memory (60%) until the end of the sessions (*Treatment*: $F_{2, 384} = 7.0$, $p < 0.001$), see Figure 1. Analysis of serial numbers of visits to the baited holes revealed that performance of Aged rats differed from those of Young and Aged + ladostigil, see Figure 2.

In the probe tests it was found, that in aged rats in the "no light" condition the presence of food did not improve RM, in contrast to Young rats. Aged rats were much more dependent on a conditional stimulus, and less able to remember the spatial position of baited holes than Young or Aged + ladostigil rats, see Figure 3.

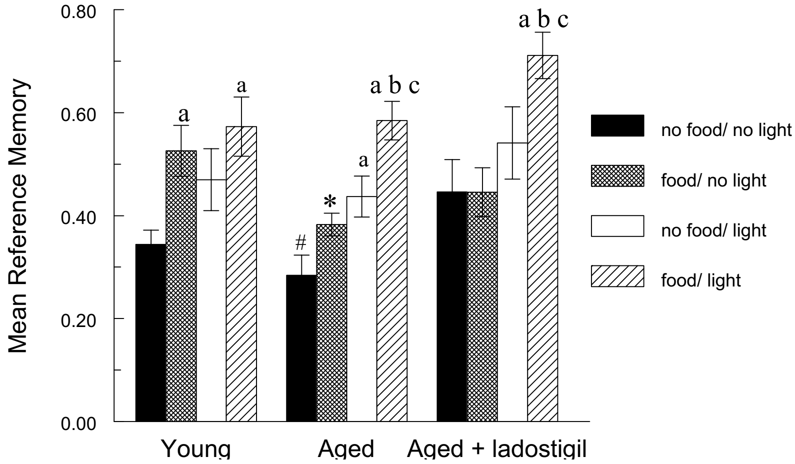


Figure 3. *Reference memory in 4 probe tests*
[a] Sig. cf "no food / no light", p<0.01; [b] sig. cf "food / no light", p<0.01 ; [c] sig. cf "no food / light", p<0.01.
**Sig diff from Young, p<0.05; [#]sig diff from Aged + ladostigil, p<0.05.*

Conclusions

The modification in the HB used in this study, the short duration and minimal food deprivation and harassment, made it sensitive enough to detect a differential contribution of natural and conditional stimuli in the expression of RM in young and older animals. Aged rats show different strategies in the performance of learning tasks characterized by a less flexibility than young rats. Aged rats have evidence of brain inflammation characterized by significant increases in activated astrocytes and microglia, which is prevented by a low dose of ladostigil (Weinstock and Shoham, unpublished observations). This may contribute towards the ability of the drug to prevent age-related cognitive deterioration.

References

1. van der Staay FJ, van Nies J, Raaijmakers W (1990). The effects of aging in rats on working and reference memory performance in a spatial holeboard discrimination task. *Behavioral and Neural Biology.* **53**, 356-370.
2. Shoham S, Bejar C, Kovalev E, Schorer-Apelbaum D, Weinstock M (2007). Ladostigil prevents gliosis, oxidative-nitrative stress and memory deficits induced by intracerebroventricular injection of streptozotocin in rats. *Neuropharmacology* **52**, 836-843.

Low-cost wind tunnel for micro-arthropods

Rostislav Zemek and Frantisek Reindl
Department of Experimental Ecology, Institute of Entomology, Biology Centre AS CR, Ceske Budejovice, Czech Republic,
rosta@entu.cas.cz

The aerial transport of tiny arthropods including wing-less species has been well documented. The initiation of aerial dispersal in most instances is not a passive phenomenon, but involves identifiable behaviors in response to specific stimuli. Here we describe a construction of wind tunnel suitable for close-up observation of micro-arthropods behavior under various conditions. The tunnel was made of glass with inner dimensions 6x6x110 centimeters. Both the inlet and outlet compartments were equipped with an electric fan, series of plastic diffusers and a metal mesh to minimize wind turbulence inside a 40 centimeters-long observation chamber in the middle of the tunnel. Wind speed is controlled by changing voltage on a variable AC/DC transformer and actual speed is monitored using a precise TESTO hot-ball anemometer. The object is observed using dissection microscope and The Observer software.

Introduction

A wind tunnel is a tool often used in chemical ecology experiments with flying insects [1-3]. The construction of such tunnel, however, does not allow a close-up observation of tiny species like *Trichogramma* spp. and other minute parasitic wasps or mites. The latter do not have wings but many are known to exhibit dispersal posture and actively take-off in the presence of wind [4]. Commercially available small-scale wind tunnels used for testing aerodynamics of three dimensional models or for calibration of anemometer probes are very expensive and their measuring chambers are usually not suitable for microscope. Therefore we designed and constructed a simple and low-cost wind tunnel which would allow to study the response of micro-arthropods to various conditions like wind speed, temperature, humidity or presence of volatiles. The example of its application in the study of *Aceria carvi* Nal. (Acari: Eriophyidae) dispersal behavior is given.

Materials and Methods

The wind tunnel is 110 centimeters long and was made of 28 glass plates glued into seven compartments with inner dimensions 6x6 centimeters (see Figure 1). The middle compartment is 40 centimeters long and is used as an observational chamber. For this purpose it has a hole drilled in the middle of bottom plate where object is inserted on a thin glass rod while it can be observed using a dissection microscope placed above the chamber. Compartments are fixed together using sticky tapes and placed on an aluminum support. This allows to take the tunnel apart when necessary (e.g. for cleaning and adding diffusers). Two 6x6 centimeters 12V electric fans fixed at inlet and outlet compartments provide source of wind. They are connected to AC/DC variable voltage transformer. The wind speed can thus be adjusted by changing voltage on the transformer. Several diffusers were tested to minimize wind turbulence inside the observation chamber. Turbulence was measured by a precise TESTO hot-ball anemometer and visualized by smoke.

The tunnel was used to study the effect of wind speed on dispersal behavior of *A. carv*i. Mites used in the experiments were collected from umbel galls sampled in caraway fields. Adult mite was placed into the tunnel and its behavior was recorded for 300 seconds by means of The Observer. The following parameters were calculated from the obtained data: (1) the time from the introduction of the mite to its first dispersal posture and (2) the time to its take-off. For each wind speed, twenty mites were examined.

Results and Conclusions

No turbulence in the observational chamber was measured when three plastic honeycomb-like diffusers and one metal mesh were inserted on both sides of the tunnel. Under this setup the wind speed could be precisely adjusted within a range between 0 and 3 m/s. Wind-tunnel experiments with *A. carvi* showed that a dispersal posture occurred at all used wind speeds but no mites took off at speed of 0.5 m/s. Speed of 1 m/s was high enough to enable aerial dispersal in some mites (see Figure 2). Time to mite dispersal behavior and take-off decreased as wind speed increased.

We can conclude that the wind tunnel described above can be build easily and with small costs while its parameters are good enough for precise control of wind speed. The tunnel is suitable for various experiments with small arthropods which need to be observed using microscope. Digital high-speed video camera mounted on the microscope will further enhance it with the possibility to capture and analyze fast movement behavior.

This work was supported by MSMT grant No. 2B06005.

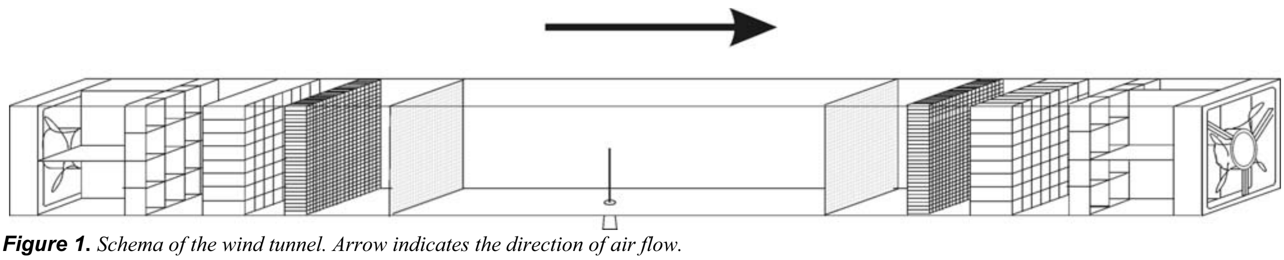


Figure 1. *Schema of the wind tunnel. Arrow indicates the direction of air flow.*

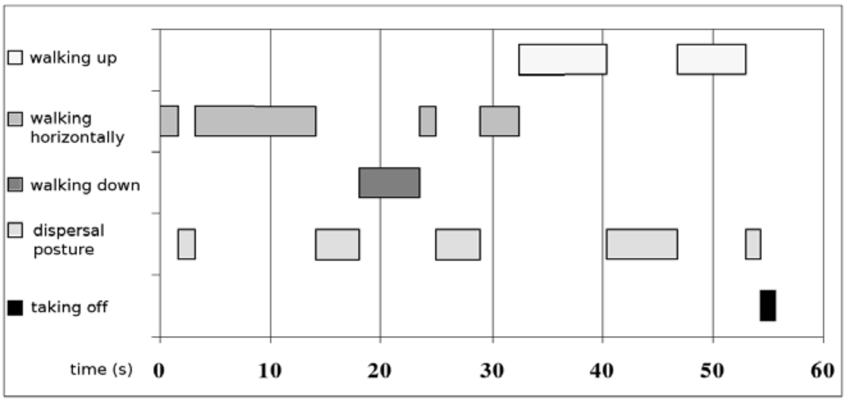


Figure 2. *Time-event plot of* Aceria carvi *behavior. The first three tracks at top indicate time when mite walked up, horizontally or down the glass rod, respectively. Dispersal posture is defined as state when mite stands up on its caudal suckers while moving its legs rapidly. The last track indicates time when mite was dislodged from the glass rod.*

References

1. Barata, E.N., Araújo, J. (2001). Olfactory orientation responses of the eucalyptus woodborer, *Phoracantha semipunctata*, to host plant in a wind tunnel. *Physiological Entomology* **26**, 26.
2. Bhasin, A., Mordue, A.J., Mordue, W. (2000). Responses of the biting midge *Culicoides impunctatus* to acetone, CO2 and 1-octen-3-ol in a wind tunnel. *Medical & Veterinary Entomology* **14**, 300.
3. McIntyre, N.E., Vaughn T.T. (1997). Effects of food deprivation and olfactory and visual cues on movement patterns of two *Eleodes* species (Coleoptera: Tenebrionidae) in a wind tunnel. *Annals of the Entomological Society of America* **90**, 260-265.
4. Smitley, D.R., Kennedy, G.G. (1985). Photo-oriented aerial-dispersal behavior of *Tetranychus urticae* (Acari: Tetranychidae) enhances escape from the leaf surface. Annals of the Entomological Society of America **78**, 609-614.

Measuring basal and complex behaviors of rats in automated social home cage systems using IntelliCage for rat technology

T. Appl[1], E. Vannoni[2], F. Buschmann[2], YK. Urbach[1], KA. Raber[1], H-P. Lipp[2,3], and S. von Hörsten[1]

[1]Section Experimental Therapy, Friedrich-Alexander University of Erlangen-Nürnberg, Erlangen, Germany, thomas.appl@ ze.uni-erlangen.de

[2]New Behavior AG, Zürich, Switzerland

[3]Institute of Anatomie, Division of Neuroanatomy & Behavior, University of Zürich, Switzerland

Introduction

Classical phenotyping of rodent models is traditionally assessed with a number of time-consuming test batteries in which animals are tested individually and experimental procedures are often difficult to standardize. For behavioral testings of mice, the IntelliCage for mice has been successfully applied for mouse phenotyping with fast and efficient test procedures for evaluation of exploratory behavior and diurnal patterns [1,2], appetive and aversive learning [3,4] and fear conditioning [5]. Based on the fact, that most of the behavioral testings have been assessed in rats, the IntelliCage for mice has been modified with the aim to test rats for their behavior within this automated home-cage system.

Aim

In order to validate the IntelliCage for rats, we introduce the transgenic huntington rat (tgHD rats) as an animal model for Huntington`s disease (HD) into an adjusted prototype of IntelliCage for rats. The tgHD rats have been shown to represent neuropathological hallmarks of HD with behavioral abnormalities including motor dysfunction, anxiety and depressive-like behavior and impaired learning abilities [6,7]. Using the IntelliCage for rats, the tgHD rats were tested for a number of behavioural parameters that have been suggested as possible read-outs for the IntelliCage, including parameters of basal activity, operant learning, spatial learning and anxiety-like behavior.

Material and Methods

At the age of six-weeks, two groups of male rats, each composed of five tgHD rats and their wildtype littermates (wt) were matched and subjected to the IntelliCage for rats. The IntelliCage for rats was designed with a technical setup composed of four conventional Typ 4 cages that are interconnected with each other, and each cage given animals access to one recording chambers (corner).

IntelliCage experiments included a contineous experimental designs including a period of habituation of animal groups to the system (seven days), nose poke adaptation (three days duration), and the learning paradigms side discrimination, reversal of side discrimination (each two days). Data were achieved by the IntelliCage software packages IntelliCage Controller, Designer and Analyzer (New Behavior AG).

Results

Within the habituation period, general behavior of the animals was monitored in the cage system with free access to water and included corner visits, nosepoke and lick performances. Comparison of the over the first 90 minutes revealed no differences between tgHD and wt rats in exploratory behavior, indicated as latency to visit the first corner and cumulative number of corner entries. In addition, we found that during habituation, the IntelliCage for rats reliably monitors total number of corner entries, nosepokes and licks in both groups, reflecting a circadian pattern with no differences in group activity. All animals visit the four corners during the first circadian interval with increased corner visits during the dark phase.

The second experimental module involved adaptation to nosepoke for fluid access and furthermore, to nosepoke on the correct side, which is indicated by a green light. Rats learned fast to nosepoke for water, reaching a stable response after 1 day of adaptation, in addition, both animals groups learned within one day to discriminate between the correct side and the incorrect side independent of the genotype (side error 22% tgHD, 19% wt). These results were re-confirmed by changing the side.

To conclude, animals accept the setup of the IntelliCage for rat, indicating that the system can be used as a suitable tool to measure rat ethological and triggered behaviors.

Discussion

Using this technical and experimental setting, additional more complex learning paradigms including place learning are currently validated in the IntelliCage for rats. Experimental modules in the IntelliCage for rats will be run successively in a longitudinal study to elaborate age-related deficits in learning and memory in the tgHD rats. Results will be compared to classical comprehensive phenotyping including tests to monitor locomotor activity, spontaneous alternation and learning paradigms in terms of reliability and sensitivity.

This work is funded by the European Union and part of the EU Research Project RATstream.

References

1. Lipp H.P., Litvin O., Galsworthy M.J., Vyssotski D.L.; Zinn P., Rau AE., Neuhäusser-Wespy F., Würbel H., Nitsch R., Wolfer D.P. (2005). Automated behavioral analysis of mice using INTELLICAGE: inter-laboratory comparisons and validation with exploratory behavior and spatial learning. Proceedings of Measuring Behavior, (Wageningen 30 August-2 September 2005):66-69.
2. Galsworthy M.J., Amrein I., Kuptsov P.A., Poletaeva I.I., Zinn P., Rau A., Vyssotski A., Lipp H.P. (2005). A comparison of wild-caught wood mice and bank voles in the Intellicage: assessing exploration, daily activity patterns and place learning paradigms. *Behavioral Brain Res* **157**:211-217.
3. Onishchenko N., Tamm C., Vather M., Hökfelt T., Johnson J.A., Johnson D.A., Ceccatelli S. (2007). Developmental exposure to Methylmercury Alters Learning and Induces Depression-like behavior in Male Mice. *Toxicological Sciences* **97(2)**: 428-437.
4. Knapska E., Walasek G., Nikolaev E., Neuhäusser-Wespy F., Lipp H.P., Kaczmarek L., Werka T. (2007). Differential involvement oft he central amygdala in appetitive versus aversive learning. *Learning and Memory* **13**:192-200.
5. Safi K., Neuhäusser-Wespy F., Welzl H., Lipp H.P. (2006). Mouse anxiety models and an example of an experimental setup using unconditioned avoidance in an automated system-Intellicage. *Cognitive, Brain, Behavior* **10 (4):** 475-488.
6. von Hörsten S., Schmitt I., Nguyen H.P., Holzmann, C., Schmidt T., Walther T., Bader M., Pabst, R., Kobbe P., Krotova J., Stiller

D., Kask A., Vaarmann A., Rathke-Hartlieb S., Schulz J.B., Grasshoff U., Bauer I., Vieira-Saecker A.M.M., Paul M., Jones L., Lindenberg K.S., Landwehrmeyer B., Bauer A., Li X-J., Riess O., (2003). Transgenic rat model of Huntington`s disease. *Human Molecular Genetics* **12(6)**: 617-624.

7. Nguyen H.P. , Koobe P., Rahne H., Wörpel T. Jäger B., Stephan M., Pabst R., Holzmann C., Riess O., Korr H., Kantor O., Petrasch Parvez E., Wetzel R., Osmand A,. von Hörsten S., (2006). Behavioral abnormalities precede neuropathological markers in rats transgenic for Huntington`s Disease. *Human Molecular Genetics*, **15 (21)**: 3177-3194.

Automatically determining active investigation in rodents using contour analysis

T.C. Lukins, M. Dewar, P.A. Crook, J.A. Heward, and J.D. Armstrong
School of Informatics, University of Edinburgh, Edinburgh, UK, {tlukins|pcrook|jheward|jda}@inf.ed.ac.uk

Introduction

The ability to automate the tracking and annotation of laboratory assays offers many possibilities for more quantitative and meaningful statistics, and could help achieve much higher levels of accuracy and constancy compared to that typically achieved across different staff and laboratories [1]. Recent advances in improving the robustness of tracking rodents using computer vision techniques [2] have lead to encouraging results, even for multiple targets. Determining, and labelling, the actions of the subject animals under observation (as for example in [3]) can provide enriched data for further analysis, and enables the potential to ultimately link behaviour back to genetic and other experimental factors.

However, one of biggest problems in handling video data of rodents is the variation with which they can deform and change direction. Correctly determining the orientation of the animal is critical to understanding their interactions with the environment and each other. In this work we address this problem by using an intrinsic description provided only by the outline of the rodent.

Method

Our approach is based on analysis of the curvature of the extracted rodent contour. We use the robust, scale invariant approach described by [4] to produce a *curvegram* profile along the perimeter. By automatically compensating for the relative complexities between different scales and shapes of contours as they change over time, the curvegram presents a unique signature in which regions of high curvature can be clearly differentiated as extrema (peaks). By taking the derivative of this profile and recording the zero crossings the number and location of these peaks can be identified.

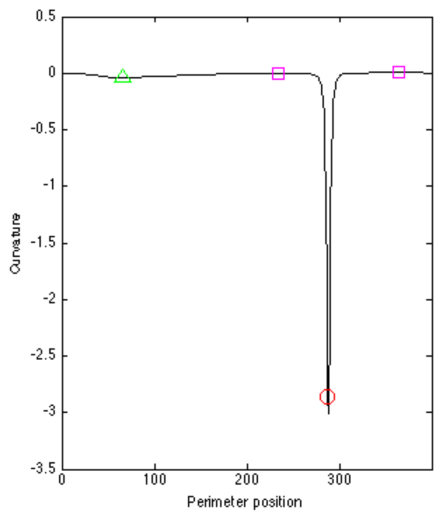
In our data, the single most detectable peak corresponds to the tip of the tail – as seen in Figure 1. Conversely, we assume the next largest peak to represent the head, or close to it. Additional peaks are often associated with the bending of the tail where it joins the body. In the case of the entire body turning around a larger number of peaks are generated, which can be used to determine this event.

Results

We generated a test sequence of 1200 frames from the top-down video of a subject animal in a plus maze. Contour data was extracted by a simple background subtraction and thresh-holding method, with additional morphological closing and cleaning. The sequence was then annotated by hand to record the location of the tail and head tip in every frame. We then ran our method over each contour frame and recorded the distance from our estimated positions – as shown in Figure 2. In addition, by recording the number of occurrences when the number of peaks in the contour exceeded four we were able to clearly distinguish the 3 incidents of the mouse turning. This is also indicated in the figure by the error in the head position at around frames 200, 750 and 1200.

Conclusions

In this work we have investigated the possibility of determining orientation from rodent contour data by considering the curvegram profile. This is robust to the appearance of additional limbs and foreshortening of the body, and is particularly able to locate the tail tip with reasonable accuracy. Further enhancements could consider other features of the contour, and using this to derive meaningful labels for the individual actions (e.g. “going forwards”, “looking right”, “rearing up”, etc).



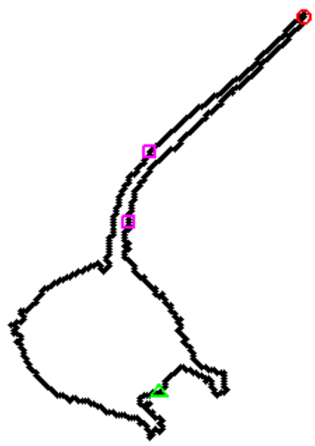


Figure 1. *Mouse extracted curvegrams and detected peak positions for tail tip (red circle), nose (green triangle), plus inner and outer bend of tail (magenta squares). Notice how the technique ignores the fore-paws.*

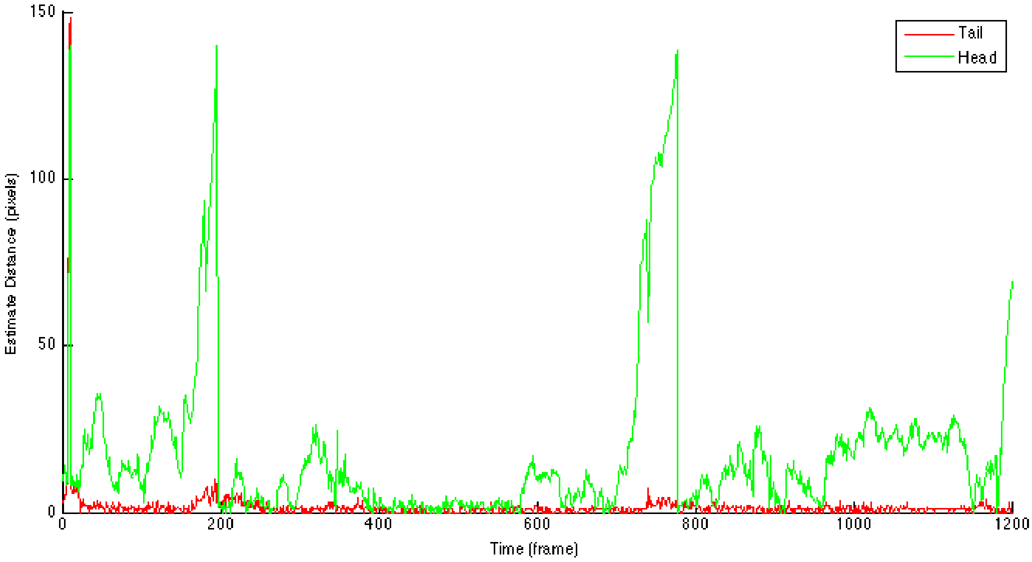


Figure 2. *Distance from estimated head and tail position from indicated location in every frame. Notice initial confusion at frame number 12 when the tail is trapped beneath the rodent.*

Acknowledgements

We thank the Genes to Cognition programme (www.genes2cognition.org) for providing the video.

References

1. Bohannon, J. (2002). Can a mouse be standardized? *Science*, **298**(5602) : 2320-2321.
2. Branson, K. and Belongie, S. (2005). Tracking multiple mouse contours (without too many samples). *In Proc. of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR'05)*.
3. Matetic, M., Ribaric, S., and Ipsic, I. (2004), Qualitative modelling and analysis of animal behaviour. *Applied Intelligence*, **21**(1) : 25-44.
4. Cesar, R.M. and da Fontoura-Costa, L. (1996). Shape characterization in natural scales by using the multiscale bending energy. *In Proc. of the 13th Int.l Conf. on Pattern Recognition (ICPR'96)*, 735-739.

Optimizing Active Avoidance Conditioning for high-throughput behavioral and cognitive screening in fish

M.Y. Ylieff[1], A. Ferrara[2], C. Fagniart[1], A. Voncken[1], and P. Poncin[1]
[1]*Behavioral Biology Unit: Ethology and Animal Psychology, Department of Environmental Sciences and Management, Faculty of Science, University of Liège, Belgium, Marc.Ylieff@ulg.ac.be*
[2]Behavioral Neuroscience and Experimental *Psychopharmacology Unit, Department of Cognitive Sciences,Faculty of Psychology and Education, University of Liège, Belgium, A.Ferrara@ulg.ac.be*

Introduction

Learning abilities have been widely documented in fishes from experimental studies. For example, a learned foraging behavior in tilapia has been characterized in food-restricted situation using methods such as “time-place learning” [1]. In more natural laboratory conditions, many behavioral or “cognitive ecology” experiments have argued strongly the social facilitation of learning devoted to locate food, avoid predator or gathering information about conspecifics (eavesdropping) [2]. Moreover, different laboratory learning paradigms typically applied in rodents have been successfully validated in fish physiology. Recently, for the first time in fish, a conditioned endocrine response (cortisol) to a stressor was demonstrated in Nile tilapia, using the Pavlovian conditioning [3].

Two-way active avoidance conditioning (“TWAAC”) is another laboratory classical technique chiefly developed for experimental psychopharmacology in rodents. The animal has to cross (at each trial) from one side to the other of a shuttle-box to avoid or escape a mild electrical shock. Thanks to the more suitable ecologically nature of the behavioral response asked to the animal, the active avoidance procedure is tested in fish for a long time [4]. Initially used to compare vertebrate learning performance, TWAAC has become more recently, a deep-rooted behavioral paradigm in goldfish. Well-known model in neuroscience research, the goldfish has permitted to elucidate some questions concerning the neurobiology of learning and memory in the context of simplified vertebrate models approach [5].

However, despite of numerous studies in teleost fishes, conditioning procedures are very few standardized. For example, TWAAC in goldfish is often time-consuming procedure (20 training days/fish) or the trial number per session can differed largely (20-40 trials) [6]. In the context of high-throughput screening for behavioral phenotyping, the optimization and the shortening of conditioning procedures in fish models are in urgent need. In this aim, we present here the major results of three TWAAC experiments in fish that manipulated different parameters: {I} Duration of the InterTrial Interval (ITI) and {II} Nature of the Unconditioned Stimulus (US) in Nile tilapia; {III} Nature of the Conditioned Stimulus (CS) in the classical “goldfish model”.

Materials and Methods

In the two first experiments, we tested a total of 2x20 Nile tilapia (*Oreochromis niloticus*) of 75-80 mm in length, randomly divided into four groups (n=10). For the last experiment, we used a total of 16 goldfish (*Carassius auratus*) of 55-70 mm in length and randomly divided into two groups (n=8).

The 56 fishes were trained and tested individually in four identical fish shuttle box active avoidance systems connected simultaneously to a Smart Control (MED associates, USA). The fish shuttle box consisted of a water-filled tank (41.0 x 20.5 x 10.5 cm) separated by a white hard plastic barrier (20.5 x 10.5 cm) into two equal compartments. There were a white light stimulus and a buzzer (CS) at each end of the tank and four stainless steel electrode plates (18.0 x 19.5 cm) at the top and the bottom of each tank. Two table top shocker modules cabinet delivered customized constant electrical voltage shocks (US). A rectangular opening (6.0 x 3.0 cm) in the barrier permitted fish to swim freely from one side of the tank to the other to escape or avoid the US. The crossing movement of the fish was monitored by four infrared light beams and their corresponding detectors located on the long sides of the tank near the rectangular opening door of the barrier. Additionally, four CCD video cameras were installed in front of each experimental tank to record the training sessions.

Each session was planned with about 20 trials of paired CS/US and the rate (%) of avoidance responses (escape during the CS presentation before the US administration) was an indicator of learning. Two-way Repeated Measures ANOVA was realized for each experiment and multiple comparisons calculated when $p \leq 0.05$.

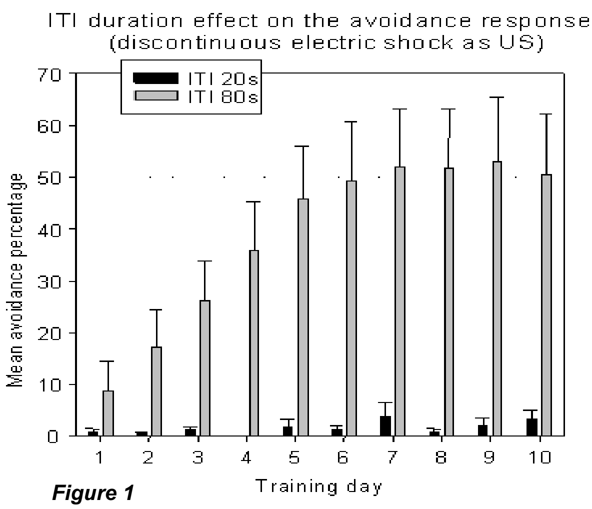


Figure 1

Results and Discussion

{I} Effects of duration of ITI on learning in Nile tilapia.

After the light (CS) was on for 15 seconds (s), a discontinuous mild electric shock (US=10V, pulsed 400 ms on and 600 ms off) was administrated, along with the light, for 15 s. In one group of fishes (n=10), the ITI was 80-s and in the second group (n=10), the ITI was 20-s. The ANOVA indicated a significant session difference [$F(9,171)=7.61$, $p<0.01$] and a strong effect of ITI on learning performances [$F(1,18)=17.54$, $p<0.01$]. The 20-s ITI group did not learn and had very low performances (see Figure 1): the post-ethographic video-analyze has shown more “like-rodent freezing behavior” in these fishes.

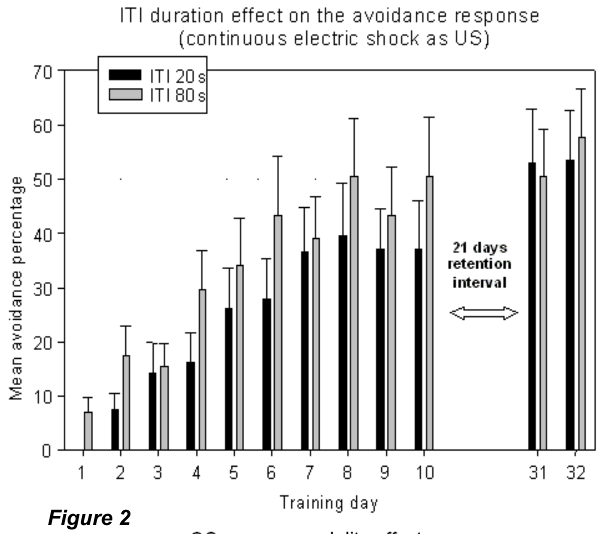


Figure 2

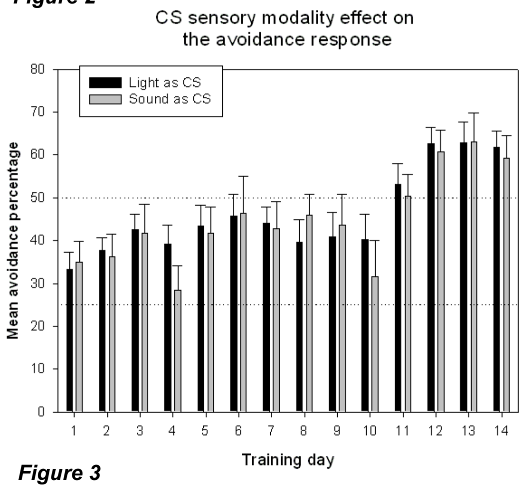


Figure 3

{II}Effects of the Nature of the US on two different ITI: impact on learning and retention in Nile tilapia.

The active avoidance paradigm was the same than in the first experiment, excepted for the US. Here, a continuous mild electric shock (10V, continuous) was administrated to the fish. The ANOVA indicated a significant session difference [$F(10,190)=14.72$, $p<0.01$] and no effect of ITI on learning performances [$F(1,18)=0.95$, $p=0.34$]. After the learning time (10 days), a test of retention was performed after 21 days. This test has demonstrated a long memorization of the task, but also a consolidation with an increase of the rate of avoidance (see Figure 2).

{III} Effects of the nature of the CS in goldfish.

Paradigm: after the CS was on for 20 s, a mild electric shock (US=5V, continuous) was administrated, along with the CS-light (group 1, n=8) or the CS-sound (group 2, n=8), for 20 s. In the two groups, the ITI was 80-s. The ANOVA indicated no difference between the groups trained with a light or a sound as CS sensory modality [$F(1,26)=0.286$, $p=0.60$]. This result would be very useful to design in the future more complex TWAAC procedures such as reversal learning. Furthermore, the pattern of learning was the same in the two groups: light or sound with 80-s ITI conducted to a slow increase of performance (see Figure 3) in goldfish.

Conclusions

This study illustrates the impact of parameters of conditioning procedures on the dynamics of learning. As a result we have found optimal combinations for high-throughput screening in main fish models: a short ITI (20 s) should be coupled with a continuous electric shock adjusted to the species (10 V in Nile tilapia, 5 V in goldfish).

References

1. Delicio, H.C., & Barreto, R.E. (2008). Time-place learning in food-restricted Nile tilapia. *Behavioural Processes* **77**, 126-130.
2. Brown, C., & Laland, K.N. (2003). Social learning in fishes: a review. *Fish and Fisheries* **4**, 280-288.
3. Moreira, P.S.A., & Volpato, G.L. (2004). Conditioning of stress in Nile tilapia. *Journal of Fish Biology* **64**, 961-969.
4. Horner, J.L., Longo, N., & Bitterman, M.E. (1961). A shuttle box for fish and a control circuit of general applicability. *American Journal of Psychology* **74**, 114-120.
5. Xu, X., Bazner, J., Qi, M., Johnson, E., & Freidhoff, R. (2003). The role of telencephalic NMDA receptors in avoidance learning in goldfish (*Carassius auratus*). *Behavioral Neuroscience* **117** (3), 548-554.
6. Portavella, M., Torres, B., & Cosme, S. (2004). Avoidance response in goldfish : emotional and temporal involvement of medial and lateral telencephalic palium. *Journal of Neuroscience* **24** (9), 2335-2342.

Effect of scopolamine on hidden- and visible-water maze learning and retention trial performance in C57BL mice

Elisabeth Piccart and Rudi D'Hooge

Laboratory of Biological Psychology, University of Leuven, Leuven, Belgium, Elisabeth.piccart@kuleuven.be

The Morris water maze is one of the most extensively used tools in behavioral neuroscience to investigate spatial learning and memory [1]. MWM learning is thought to rely extensively on hippocampus, and involves several major neurotranmitter systems. One such major neurotransmitter system of great importance is the cholinergic system [2]. Scopolamine, a muscarinic cholinergic antagonist, is known to cause impairments in MWM testing [1]. Studies have suggested that scopolamine administration specifically impairs the development of spatial navigation strategies, thus the acquisition, rather than memory consolidation or recall [2,5]. Day and Schallert hypothesized that acquisition impairment, following anticholinergic treatmentis mediated at least partially by entrapment in an inefficient, non-place strategies [3]. The present study investigates the effect of scopolamine in MWM. Subjects were trained during 10 trialblocks, each of which consisted out of 4 trials. After the fifth and the tenth training trial, animals were tested (probe trial of 100s). Animals (n=4) were injected with scopolamine (1mg/kg), 30 minutes prior to performance. Animals (n=4) injected with saline (10ml/kg) served as controls. To examine whether the impairment is due to deficits in acquisition or recall, we injected control animals (n=2) with scopolamine (1mg/kg) 30 minutes before the second probe trial. In another trialblock all animals (n=8) received saline injections (10ml/kg).

Scopolamine injection during acquisition affected MWM learning as well as retention trial performance (see figure 1) (effect of treatment: $p<.001$; effect of trialblock: $p<.001$; interaction: $p=.07$). Considering the learning curve, the impairment of scopolamine treated animals seemed most pronounced in the first week of training, after which constancy is reached within both groups. On both probe trials, we do not find significant differences between scopolamine treated animals and controls (Probe 1: effect of treatment: $p=.172$; effect of quadrant: $p<.001$, interaction: $p=.003$; Probe 2: effect of treatment: $p=1.000$, effect of quadrant: $p=.08$, interaction: $p=.813$).

Control animals that received scopolamine before the second probe trial did not perform differently from saline-treated animals ($p=.514$) (see Figure 2). We conclude that scopolamine affected acquisition performance but not performance on probe trials. It could be argued that indirect effects of scopolamine could underlie the effects of this drug (e.g.,hyperactivity, resulting from scopolamine administration, can make it more difficult for the animals to stay on the platform). In a final experiment, we assessed the effects of scopolamine treatment on visible-platform MWM performance. No difference in latency between scopolamine-treated animals and controls was found in the visible-platform condition ($p = .486$), which suggests that motivation, nor locomotor hyperactivity underly the differences found during MWM acquisition in laboratory mice. Thus, the present findings indicate that scopolamine affects central brain mechanisms underlying spatial learning, possibly in relation to previously reported impairment of inhibitory avoidance behavior following disturbed amygdaloid cholinergic functions [5].

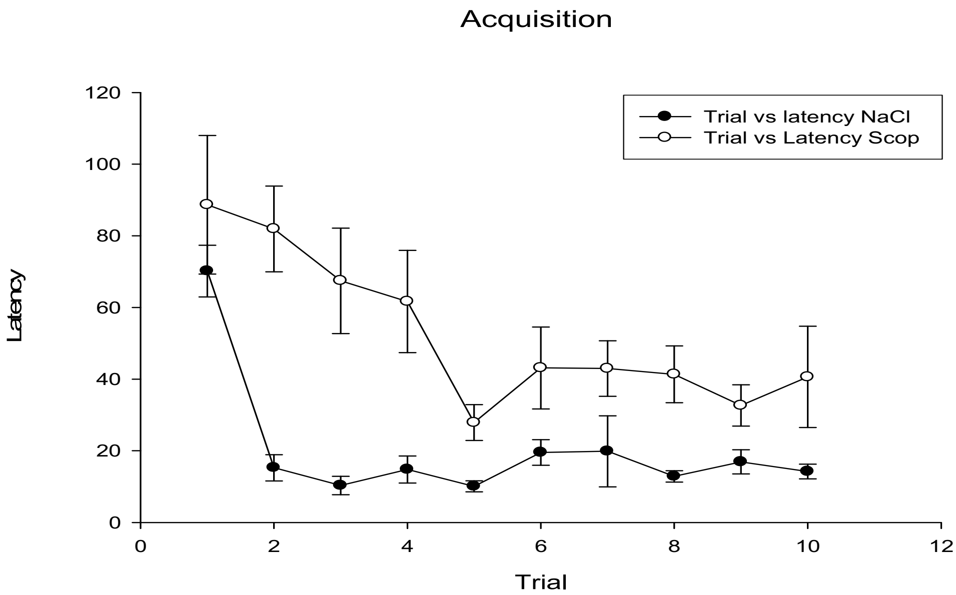


Figure 1. *The effect of scopolamine on the acquisition of the Morris water maze.*

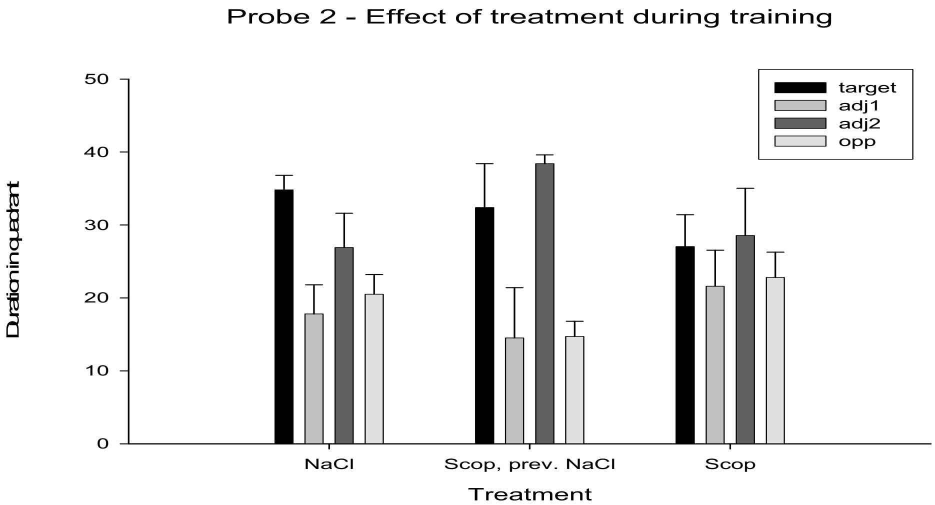


Figure 2. *Effect of previous treatment during the second probe trial.*

References

1. Hchoi, S.H. Woodlee, M.T., Hong, J.J., Schallert, T. (2006). A simple modification of the water maze to enhance daily detection of spatial memory in rats and mice. *Journal of Neuroscience Methods* **156**, 182-193.
2. D'Hooge, R. De Deyn, P.P., (2001). Applications of the Morris water maze in the study of learning and memory. *Brain Research Reviews* **36**, 60-90.
3. Day, L.B., Schallert, T. (1996). Anticholinergic effects on acquisition of place learning in the Morris Water Task: Spatial mapping deficit or inability to inhibit nonplace strategies. *Behavioral Neuroscience* **110**, 998-1005.
4. Baldi, E., Lorenzini, C. A., Bucherelli, C. (2003). Task solving by procedural strategies in the Morris water maze. *Physiology & Behavior* **78,** 785-793.
5. Riekkinen Jr., P., Kuitunen, J., Riekkinen, M. (1995). Effects of scopolamine into the anterior and posterior cingulate on passive avoidance and water maze navigation. *Brain Research* **685**, 46-54.

A system for measuring motor abilities in adult mice in neurological experiments

Magdalena Mazur[1] and Bartosz Karaszewski[2]
[1]*Department of Biomedical Engineering, Gdansk University of Technology, Gdansk, Poland, magda@biomed.eti.pg.gda.pl*
[2]*Department of Neurology of Adults, Medical University of Gdansk, Gdansk, Poland*

The rapid technology development caused that machines took over some part of humans' work. The field of computer analysis of experiments made on animals is intensively evaluated. Automatic system analyzing animals' movement is much more precise and objective than human, it is also devoid of susceptibility to optical illusions. Some values, velocity for example, are also difficult to estimate by human while computer programs have no problems with precise calculating this value. Another problem concerns long-lasting experiments. People are not able to stay focus for more than few hours, consciousness of the computer lasts as long as it is necessary. All these issues were reason for creating an automatic system for recording and analyzing motor abilities of mice with brain focal lesion. We needed a system that can precisely measure different aspects of mouse velocity, works at least 24 hours and monitors 3 mice simultaneously. The System for Analyzing Neurological Experiments (SABN) was created.

The system consists of a maze made from plexiglass, a video camera and a software for recording, analyzing and storing data. The camcorder is situated above the maze and gives a top view. Required picture definition is 320 x 240 pixels. The maze consists of three identical, separate labyrinths, see Figure 1. It is 36 cm long, 53 cm wide and 5 cm high. Floor and walls between each labyrinth are painted white to give indispensable contrast to black mice and to prevent them from seeing each other. Inside the maze there are obstructions like steps, ramp, direction choice branching and curves. Steps and ramp are made for dexterity observation, direction choice branching – for investigating hemiplegia. Water and food containers were placed in different parts of the labyrinth to enforce animals to overcome the obstacles and run around its corridors. System is using USB interface to communicate with the camcorder. Created software consists of:

- program for recording experiments,
- program for analyzing recorded video,
- database with interface.

The SABN was written in C++ and MySQL. The relationships between all parts of the software can be seen in Figure 2.

The program for recording experiments communicates with the camcorder and records compressed frames. The results of recording program (set of images) are analyzed after change into binary image, finding the characteristic points and marking the field of analysis, see Figure 3. System works on the basis of color and size distinction. It is insusceptible for other dark objects that occur in the maze during experiments (excrements, shadows). Nevertheless, when external disturbances appear, system may not work properly. The centers of gravity of all frames illustrate the general draft of a rodent motion in the maze.

When the last picture is analyzed, program counts parameters of the experiment:

- mean velocity [cm/s],
- velocity during movement [cm/s],
- time of rest (percentage of total time in the maze [%]),
- velocity at the steps [cm/s],
- velocity at the ramp [cm/s],
- time spent by the water dispenser [%],
- number of each side chosen in the branching.

Velocity is counted from the mouse displacement between two frames (frequency of frame capturing is known). Mean velocity is a total velocity within all 24 hours of experiment. Velocity of the movement is estimated during the time when motion is observed by the system. It does not include time of rest, grooming or eating. Velocity at the steps and ramp is space-limited to the area where obstructions are placed.

The results of all experiments are stored in the database and ready for further analysis.

There is also a possibility to extend system with an additional side camcorder and a 2 cm high bar with 1cm x 1cm holes. This extension enables observation of agility by counting the number of cases where mouse's paw falls into the hole. The system is counting motion in area under the bar.

Authors do not exclude creating other extensions for better diagnosis of physical degenerations induced by brain damages.

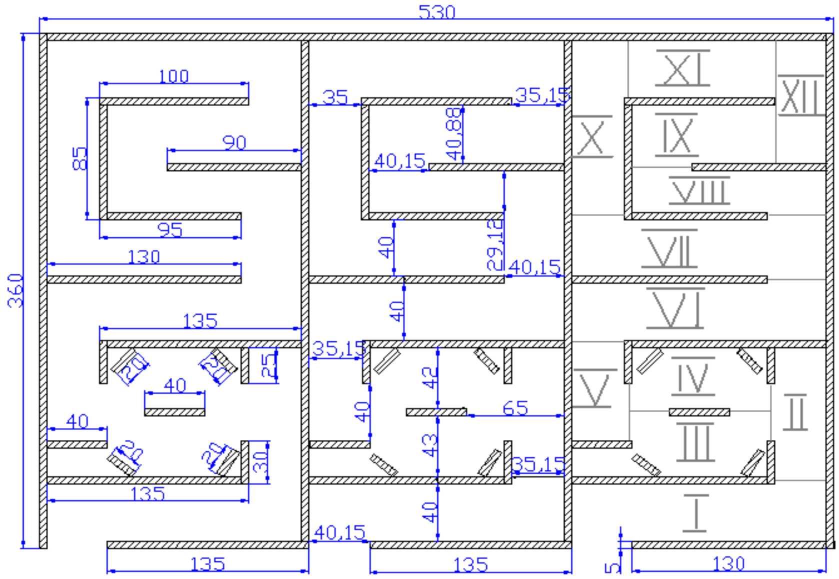


Figure. 1 *A schema of the maze*

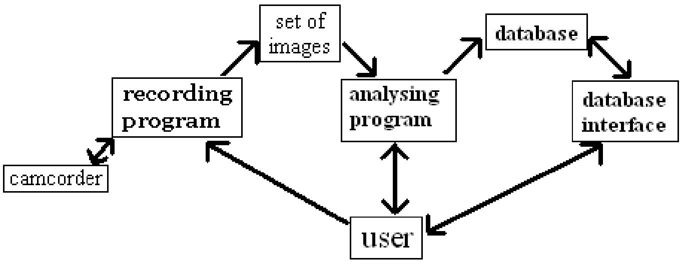


Figure 2. *A schema of the software*

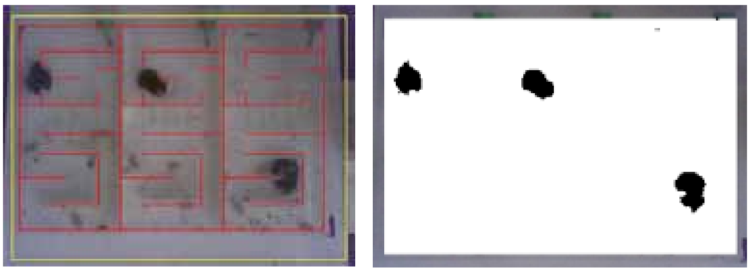


Figure 3. *Image of marked field of analysis and after thresholding*

Fire safety architecture - measuring the process of orientation

M.D. Bredero
owner office Architect Maarten Douwe Bredero and a lecturer at Saxion Universities
Kortestraat 27, 7419 CK Deventer, The Netherlands, mail@architectmaartendouwebredero.com

Key words
Architecture, Fire Safety Engineering, Evacuation, Clearance time, Self-rescue, Orientation, Perception, Space

Introduction

The problem is that evacuation during instances of fire takes too much time, especially in high-rise buildings. Several reasons lie behind this phenomenon [1]. Wtihin the domain of Self-rescue, this experiment solely involves the aspect of orientation.

A fictional design (see figure 1) for a high-rise office building and museum assumes fire escape optimisation with a widening of the route along the facade or widening within the core of a tower.

The corresponding hypothesis states that, regardless of their projection, escape routes which widen while going down considerably shorten the perilous period of clearance.

In addition, the two fundamentally different ways of circulation (entrance) in a given volume i.e. building envelop are to be compared [2].

Questions of research

Of interest are:

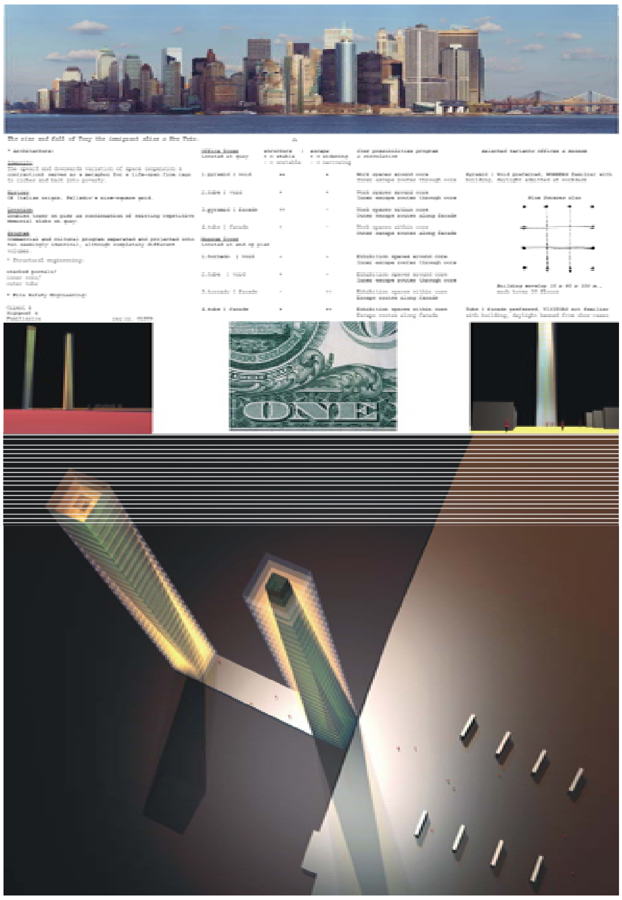


***Figure 1**. Architectural model of the two towers, 10x10x100m. each.*

Do widening escape routes indeed shorten the evacuation time? Which flight goes faster, a flight along a façade with sight on the surroundings? Or in fact a flight through an atrium while having visually contact with other persons (part of the crowd) trying to escape?

Other issues to be researched will be:

Does fleeing through relatively high, shallow, curved and bright spaces go quicker? Is flight along relatively 'warmly decorated', light colored and soberly furnished routes more efficient [3]?

Sort of research

A two-folded approach implies the following:

The intention is empirical research with an inductive character. From specific observations of behavior an attempt is made, on the one hand, to deduce general rules. On the other hand it is presumed, as a deductive hypothesis, that possibilities for orientation indeed shorten the time of escape during the hazard of fire.

Given the fact of applied and not purely scientific research, it will not be like a survey. Instead the study will be characterized as explorative (case study) and also as testing (hypothesis). Depending on the number of measurements, in first instance the research will be qualitative in nature. Therefore only limited statistical calculations will be necessary. Registration of human behaviour during the evacuation process of way-finding will provide an insight into the psychological perception of spatial information.

The objective naturally is a neutral result, due to the application of a fitting methodology and adequate focus [4].

Method and implementation of research

Precise measurement is achieved as follows:

Behavior of PC-operating, serious gaming, persons is monitored while they run through the Virtual Design looking for a way out (filming eye movements and changes in bodily functions such as heart beat). The actual path taken through the various spaces in the model is tracked as well. A shorter way-finding time implies a shorter clearance time in reality.

The variable of research is hence flight-duration, for which a clock is observed. The trustworthiness of the measurement is easier to deal with then its validity.

Moreover, relevant questions regarding orientation i.e. the mental condition of subjects before and after the tests will be put forward and compared with the actual outcome.

The present architectural sketch will be transformed into a Preliminary Design which hence fort will be converted into a semi realistic environment using advanced visualization and gaming techniques.

Subjects will explore the virtual interior and exterior through computer interfaces. They will interact with computer screens, sitting on stools behind gaming consoles and shall be wearing special glasses.

Via this non-direct method (registration of behavior) the mental process of orientation is measured (see figure 2).

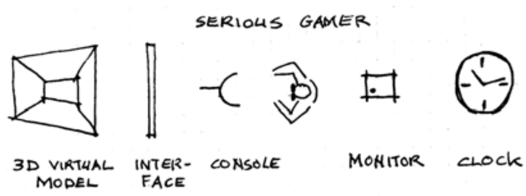


Figure 2*. Set up of measuring the interaction with virtual environment.*

General constraints

Realistic model (Architectural Preliminary or Definite Design transformed into real-life visualization);

Representative measurement (large enough number of subjects);

Applicable (poper questionaire and usable protocol for professional architects as result).

References

1. Kobes, M. Zelfredzaamheid bij Brand. Tien mythen ontkracht. NIFV-NIBRA, to be published shortly.
2. Garling, T., Book, A., Lindberg, E., (1986) Spatial Orientation and Wayfinding in the Designed Environment. Journal *of Architecture, Plans and Resoursces* **3**, 55-64;
3. Passini, R., (1992) Wayfinding in Architecture. Van Nostrand Reinhold, New York, second edition;
4. Werner S., Schindler, L.E., (2004) The Role of Spatial Reference Frames in Architecture. *Environment and Behavior* **36(4)**, 461-482.

Development of a software tool for correlating user experience to network QoS in a mobile network

T. Deryckere[1], W. Joseph[1], L. Martens[1], L. De Marez[2], and K. De Moor[2]
[1]*Department of Information Technology, Ghent University/IBBT, Gaston Crommenlaan 8, B-9050 Ghent, Belgium*
{tom.deryckere, wout.joseph, luc.martens}@intec.ugent.be
[2]*Department of Communication Sciences, Ghent University/IBBT, Korte Meer 7-9-11, B-9000 Ghent, Belgium*
{Lieven.Demarez, KatrienR.DeMoor}@UGent.be

Introduction

In this era of converging media technologies, value is increasingly migrating from products to experiences as customers seek out truly personalized value to satisfy their situational needs. It is often hard to measure and to predict what the user experience will be during service consumption. This is nevertheless a very important aspect that should be taken into account while developing applications or frameworks.

Hence, the ultimate measure for mobile media networks and services is how the end-user perceives the quality of the new media and services. Measuring end-user perception is very complex given the number of parameters (context, network, activity, device) that can influence this perception. This paper demonstrates the preliminary result of a monitoring probe and interdisciplinary methodology that will capture user experience and correlate this to the wireless reception quality with the goal of quantifying and predicting user experience.

Methodology

QoE dimensions

User experience (QoE) is a multidimensional concept that consists of both social and technical axes [**Error! Reference source not found.**]. User experience is not simply related to technical excellence: "*It is possible to have excellent Quality of Service (QoS) and poor QoE*" [**Error! Reference source not found.**]. Parameters and influences come from different layers and have to be measured in appropriate ways. This includes a multidisciplinary approach where not only objective parameters such as network quality and device capabilities are relevant. Also the more subjective parameters like expectations, emotions, usability, and context must be taken into account.

In [**Error! Reference source not found.**] we defined QoE dimensions in 5 building blocks. These blocks structure the different aspects influencing the user experience: *Quality of Effectiveness* (is the application, network, or device doing what it is supposed to do?), *Quality of Efficiency* (does the application or device work well enough for the user?), *usability* (deals with how easy it is for the user to accomplish tasks), *expectations* (the QoE will be influenced by the degree to which the expectations of the user are met), *context* (different contexts exist that can influence the experience: the environment, the social context, cultural context).

Probe model for data collection on a multidisciplinary level

In order to collect the relevant data, a probe model that measures data across the different dimensions influencing the user experience is created. The concept of this model is shown in Figure 1. Since the dimensions reflect different aspects of classical research domains (technological and social research) we have created a software model consisting of three layers. Each layer consists of one or more software monitoring probes (Figure 1). Each probe fulfils a specific task:

The contextual probes consist of software probes that deal with determining what the context of the application usage is. This can exist of GPS data (environmental context), information coming from the users agenda, or data reflecting the users' mood or activities.

The experience probes consist of the software probes having build-in intelligence in order to capture the user experience. To this end, automatic questionnaires completed by the user on the mobile device before, after, or even during application usage could be a possible mechanism. Other ways are detecting application usage by monitoring keystrokes.

The QoS probes consist of the software probes that will deal with the monitoring of the technical parameters such as network performance (throughput, delay, signal strength), device performance and capabilities (memory usage, screen size), and application properties (video codec).

Partitioning of the monitoring model in these three layers enables collaboration with experts with different backgrounds such as social researchers, engineers, and usability designers.

We have developed a software tool that reflects the probe model. The implementation of the client software was done in C# within the .NET Compact Framework 2.0 and by using Windows Forms. Auxiliary classes were taken from the Smart Device Framework v2.1 from OpenNetCF. For each category of probes, new modules can be created that can reflect new parameters to be monitored during service consumption. This software tool runs during the application usage and is

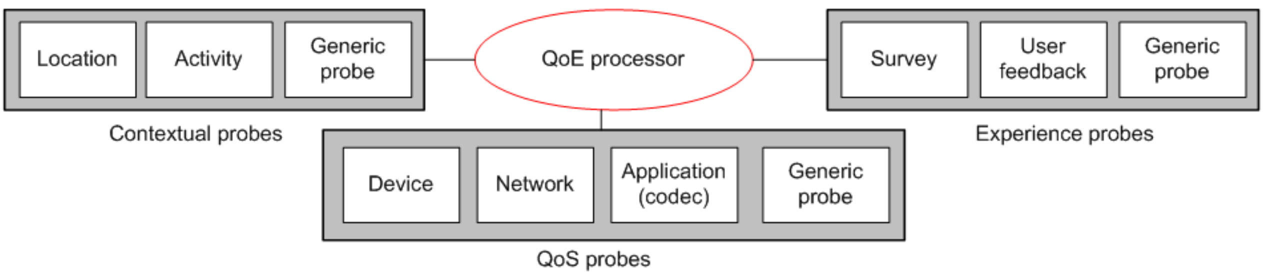


***Figure 1**. Structure of a software model for montoring QoE.*

connected to a backend infrastructure that stores and analyses the data.

User tests

It is of key importance to be able to evaluate the user experience with a large test panel. For this test panel we will use the i-City Living Lab environment (http://www.i-city.be). I-City is a living lab in Belgium that exists of a city-wide wireless WiFi infrastructure with more than 1100 test users. For this we have integrated a software tool (Section II.B) into the i-City service platform. We have also chosen to target a specific application that currently runs on the I-City platform that will be monitored. Users are given a Personal Digital Assisent (PDA) of type HP IPAQ rw 6815. The application that is tested is a Wineguide that assists people in searching and finding information about wines. The application also gives the possibility to create a personal wine collection.

After selection of the application we defined different usage scenarios that had to be completed by the test users. 10 test-users completed the scenarios under different reception levels. During usage of the application we monitored the signal strength (calculated from the Received Signal Strength Indication RSSI). After application usage a short experience-questionnaire of 6 questions was presented to the users on the PDA. In short, the users were asked to report their experiences with the Wineguide application on several dimensions by means of 5-point Likert scales. The data are collected in order to correlate the answers with the reception quality.

Results and conclusions

The method to evaluate user experience needs a lot of data entry points and a multidisciplinary approach. By defining Quality of Experience dimensions and the creation of a multilayered probe model we are able to study the correlation between the different dimensions. In addition, the proposed method integrates a number of fragmented traditions and findings into one meaningful, multidisciplinary approach. Furthermore this system allows researchers from different disciplines to quantify these dimensions of the user experience. Results obtained using the described methodology will be presented at the conference.

Acknowledgments

This work was supported by the IBBT – ROMAS (Research On Mobile Applications & Services) project, co-funded by the IBBT (Interdisciplinary institute for BroadBand Technology), a research institute founded by the Flemish Government in 2004, and the involved companies and institutions. W. Joseph is a post-doctoral fellow of the FWO-V (Research foundation-Flanders).

References

1. Perkis, A., Munkeby, S., and Hillestad, O. I. (2006). A model for measuring quality of experience. *Proceedings of the 7th Nordic Signal Processing Symposium, 2006. NORSIG 2006.* 198-201.
2. Empirix. (2003). Assuring QoE on Next Generation Networks - White Paper [online] Available from: http://www.voip-report.com/wp_download/whitepaper_NGNT_AssuringQoE.pdf [Accessed 22 Feb 2008].
3. K. De Moor, L. De Marez (2007), Quality of (user) Experience: conceptualization and measurement of rising concept., *IAMCR Conference (Paris 2007),* [CD-ROM] 1-14.

Predicting group performance using group indices: A new proposal

A. Andrés, Ll. Salafranca, and A. Solanas

Deparment of Behavioral Sciences Methods, University of Barcelona, Barcelona, Spain, aandres@ub.edu

This research is focused on the influence of different psychological variables over group performance. Specifically, personality measures, the Skew-symmetry index Φ [1], and generalized/dyadic measurements of reciprocity of the Social Relations Model [2] were analyzed. The objective was to examine whether personality measures, Φ values or generalize/dyadic (or both) reciprocity measurements in interpersonal perceptions could explain group performance (following a round robin paradigm).

Method

In order to carry out the study, a sample of 64 participants was gathered forming 16 groups of four people. Each group was enrolled in a laboratory task consisting of two sessions of 60 minutes each. During the first session, they were given the "Foundation Task" [3] to get accustomed to the methodology of the task and they answered the personality questionnaire NEO-FFI [4]. In the second session, the instruction was to reach as many agreements as they could in a set of problem solving tasks. Due to the characteristics of this kind of tasks (without a unique valid solution) participants were forced to interact during the session to reach solutions. Interaction allowed them to form interpersonal perceptions about how members of the group collaborated to achieve the target. At the end of this session, participants filed in a questionnaire in which they had to rate each partner (round robin rating). The scores reflected participants' perception about the contribution of their partners to reach the maximum number of agreements. The Skew symmetry index and general/dyadic reciprocity measures were calculated from the answers to this questionnaire. The number of agreements in the second session was taken as an indicator of group performance. A discrepancy index between the scores in NEO-FFI was developed and computed as an index of groups' personality.

Results

Our results show moderate negative correlations between the number of agreements and phi value of item 1 (*She/He profits from time to solve the task*), generalized reciprocity of interpersonal perceptions of item 6 (*Her/His dialog result helpful to solve the task*), and discrepancy index of Neuroticism, r= -.560 (p= .24), r= -.504 (p=.047) and r= -.555 (p=.026), respectively. The regression analysis shows that the Φ value for item 1, generalized reciprocity measure of item 6 and personality discrepancy in Neuroticism explain together 70.4% (p = .002) of groups performance variance. Φ value of item 1 accounts for 31.4% of this variance and including generalized reciprocity measure of item 6 and discrepancy in Neuroticism this percentage increases 25.7% and 13.3% respectively.

Conclusions

These results partially support that group measures can be useful to predict groups' performance in social environments. In our research, results showed that a discrepancy in perceptions relating to a useful dialog, the profit of time and scores in Neuroticism is related to the number of agreements. These results show that group measures predict more percentage of variance than those reported by some authors [5,6] with predictions between -.06 and .22.

References

1. Solanas, A., Salafranca, Ll., Riba, C., Sierra, V. and Leiva, D. (2006). Quantifying social asymmetric structures. *Behavior Research Methods*, ***38***, 390-399.
2. Kenny, D. A. (1994). *Interpersonal Perception. A social Relations Analysis.* The Guilford Press: New York.
3. Watson, R. T. (1987). A study of group decision support system use in three and four-person groups for a preference allocation decision. Unpublished Ph.D. dissertation, University of Minnesota, Minneapolis.
4. Costa, P. T., Jr., & McRae, R. R. (2002). *Manual NEO PI-R.* Madrid: Tea Ediciones.
5. Hough, L. M. (1992). The 'big five' personality variables--construct confusion: Description versus prediction. *Human Performance,* ***5***, 139.
6. Mount, M. K.; Murray, R. B. and Stewart, G. L. (1998). Five factor model of personality and performance in jobs involving interpersonal interactions. *Human performance*, **11**, 145-165.

Pain structuring in brain codes

I.K. Shekhovtsev
Department of Psychophysiology, State University-High School Economics, Moscow, eshekhovtsev@hse.ru

Summary

The electroencephalographic method is charged with developing practice parameters for experts for diagnostic procedures, care modalities, and former disorders. The selection of topics for which practice parameters are developed is based on prevalence, frequency of use, economic impact, membership involvement, controversy, urgency, external constraints, and resources required.

All adults have low back pain at some time during their lives. About 14% have an episode of back pain that persist longer than two weeks, but only 2% have back pain associated with spinal stenosis [1]. Self-esteems one of the basic diagnostic tool for pain evaluation [8]. As an excellent review by Kahneman points out, systematic efforts to organize the self-esteem of personality began shortly after Digman suggestion, although such efforts appear to be more surely linked to psychologists, than to Caettel suggested that a careful analysis of would assist the understanding of personality, and this stimulated John and Pollck to examine personality terms commonly found in the EEG. For the above-stated reasons, it would be of great importance to find a biological marker that could help psychologists in making a differential diagnosis and selecting performances [2, 4, 5, 9].

Method

The EEG and also the set of self-esteems reports were recorded simultaneously from 37 patients. The 47 various types of behavior reflected in the consequences of the personality (the will, the resoluteness, the perseverance, the politeness, the pain, etc.) according to scale from 0 to 5 evolve in relationship to the personal considerable body of 300 EEG brain codes extension were investigated under linear regression model with the use of a digital computer.

Results

A comparison between the physiological and the psychological regularities reveals the number of diagnostics equations with four arguments respectively.

One example are shown electrophysiological correlates (R=0.54) in which the EEG activity represented in low back pain (M54.5 in ICD-10). Modern mamagement of the low back pain (LBP) may have altered the syndrome itself. Proiferating diagnostic modalities, activities of nonmedical practitioners, and the growth of the disability compensation systems may have become determinants in the natural history of the condition. Three first predictors combined 92% of the diagnostic significance. First include 32% in configuration of a slow 5 Hz waves. This generalized theta excess has been reported to be accompanied by alpha [6]. Second and third related to the characteristics of the coherent oscillations complexes and covered 60% of diagnostics. And anticipated upon the assumption that cognitive space is coded by groups of cyclically recurrent neural impulses which are generated by coherent discharges of central neurons. Chains of these code elements are proposed to form code words. It is asserted that code words are generated by ensembles of neurons located in different brain structures and that they can recur cyclically with oscillations equal to alpha [7]. Four predictor under diagnostics weight 8% associated with beta 17 Hz oscillations in left central brain area.

It is clear that biological markers related identification between individuals. Some comparisons with predictors of preferences indicated that pain manifistation seem to be strongly associated with the condition work hard in business.

Discussion

Investigators who rely on self-report data in health surveys, epidemiologic studies, and clinical trials assume that the respondent understands the questions and terminology in the same way that the investigator does, accurately recalls information, and accurately formulates answers. These assumptions may or may not be accurate or reasonable []. These questions place a large cognitive burden on the ability of respondents to understand what is being asked and to report events from their lives, based on a system of memory [3]. A number of variables affect the ability of respondents/patients to accurately report information. Some of these variables include instructions, respondent-selected response strategies, the period of time that the information has been held in memory, the period of time to which the question refers, the format of the response alternatives, the order of the questions, and the affect of the respondent. For reviews of this literature, see [10]. The pain remains should not to be determined at improve to cognitive space with respect to all others prefrences like endurance; frankness, truthfulness; diligence; courage; accuracy; friendliness; etc. Pavlov says about choosing best experimental constructs only if they developed by life.

Conclusion

The different meanings reached by studies investigating the utility of EEG prevent a confident conclusion regarding utility. Some investigators are strong believers in the diagnostic usefulness of this technique. Others are more skeptical. One possible cause is variable study design. Most studies provide evidence no higher than expert opinions or case report. Controlled studies, especially with blinded interpreters of the tests, would provide needed evidence of efficacy in view of the existing confusion in the literature. It is our consensus that the current evidence supporting the diagnostic use of EEG is primary contact tool and that EEG should be regarded as satisfactory, meaning that current evidence is insufficient to determine appropriateness. This group does encourage further study. In this paper, we summarize the major findings of neurobiology studies on self-esteems, highlighting convergent points of view about physiological substrates. Consistent with these investigations, we describe our results using the multiple linear regression method in an exploratory fashion in the brain codes defined subtypes of satisfaction oneself obtained with EEG analysis. We also examine the diagnostic role of these techniques in the evaluation others self-concepts. According to our findings, we propose a brain code model for prognostics that substantially involves the behavior pathways.

References

1. Deyo R.A. Early diagnostic evaluation of low back pain. (1986) J Gen Intern Med **1**: 328-338
2. Digman J.M., (1990) Personality structure: emergence of the five-factor model. Annu. Rev. Psychol. **41**, 417-440
3. Jobe J.B.Cognitive psychology and self-reports: Models and methods. (2003) Quality of Life Research **12**: 219–227.
4. John E.R., Prichep L.S., Friedman J., Easton P. (1988). Neurometrics: computer-assisted differential diagnosis of brain dysfunctions. Science **293**,162-169.
5. Kahneman D., Krueger A.B., Schkade D., et al. (2006) Would You Be Happier If You Were Richer? A Focusing Illusion. Science **312,** 1908-1910
6. Klimesch W., (1999) EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. Brain Res. Brain Res. Rev. **29**, 169-195
7. Lebedev A.N., (2001) The oscillatory mechanisms of memory. Cognitive Processing **2,** 57-66
8. Nabe-Nielsen K, Fallentin N, Christensen KB, Jensen JN, Diderichsen F.Comparison of two self-reported measures of physical work demands in hospital personnel: a cross-sectional study.(2008) BMC Musculoskelet Disord. 2008 Apr **29**;9:61.
9. Pollock V.E., Schneider L.S. (1990). Quantitative, waking EEG research on depression. Biol Psychiatry **27**, 757-780.
10. Tourangeau R, Rips LJ, Rasinski K. The Psychology of Survey Response. New York: Cambridge University Press, 2000.

The Measurement of Thai Guidance Teacher Professional Identity and Role Performance

S. Seedagulrit

Department of Psychology and Guidance, Loei Rajabhat University, Loei, Thailand, somsaksee@hotmail.com

The purpose of this paper was to measuring professional identity and role performance of Thai guidance teacher (TGT) by examining the fieldwork experience through the eyes of TGT themselves. Conceptualized from a symbolic interaction framework and Stryker,s identity theory. Summated rating scales (ranging from 1 to 5) form the measure of TGT professional identity and role performance. This study used the Critical Incidents Technique (CIT)[1] to gather information on the events and interactions which TGT identified as critical in their professional development. Five procedural steps for CIT include: 1) identifying the general aims of the activity; 2) establishing precise instructions for observing the activity; 3) collecting the data; 4) analyzing the data; and 5) interpreting and reporting the data. Participants were 26 secondary school guidance teachers. Participants were asked to describe both a positive and a negative event or interaction which occurred during their fieldwork experience and which contributed to their sense of professional identity and role performance. A total of 165 critical incidents for professional identities and 160 critical incidents for role performances were submitted.

The critical incidents were categorized into 3 themes for professional identity and 4 themes for role behaviors. In order of frequency, the themes of professional identities were: Personality Prominence (N = 76), professional self-esteem (N=60) and Identity Salience (N= 29), the themes of role performances were: Guidance and Counseling Service (N= 65), Guidance Administration and Management (N= 40), Co-ordination (N= 35) and Research and Development (N= 20).

The seven themes were further divided to summated 5-point rating scales. After administering the survey, factor analyses using maximum likelihood extraction and oblique rotation were undertaken.

The data collected from the 102 TGT pilots. This iterative process result in 29 items measuring 2 of the 3 hypothesized factors of identity scales and 48 items measuring 4 of the 4 hypothesized factors of performance scales being retained. The final copy of the survey was analyzed to gain an indication of its reliability, with the aim being to measure the internal consistency. The reliability coefficient on this final form of the professional identity survey yielded an alpha of .99 and of the role performance survey yielded an alpha of .99. The reliability coefficients for the scales were:

TGT Professional Identity

1. Identity Salience and Personality Prominence	.9909
2. Professional self-esteem	.9135

TGT Role Performance

1. Guidance Administration and Management	.9752
2. Guidance and Counseling Service	.9645
3. Co-ordination	.9517
4. Research and Development	.9275

The measurement of TGT professional identity and role performance by examining the fieldwork experience were found to be significant factor to describe the role behavior of Thai guidance teacher and useful to measuring the professional socialization of new guidance teacher increasing sense of professional identity and role performance.

References

1. Flanagan, J.C.(1954). The critical incident technique. *Psychological Bulletin* **51**, 327-358.

A simple and effective method for measuring brain asymmetry in children with autism

R. A. Wittling[1], E. Schweiger[1], L.Yu. Rizhova[1], E.A. Vershinina[2], and L.B. Starup[2]

[1]FinTec GmbH, Neunkirchen FRG, Center for Neuopsychologial Research, University of Trier, Trier FRG; [2]Autisme Centre Vestsjelland, Slagelse, Denmark; Pavlov Institute of Physiology of Russian Academy of Sciences, St. Petersburg, Russia

arne-wittling@znf.uni-trier.de

For the last two decades a number of publications have shown that the pattern of brain asymmetry in people with mental handicap, including autism, depression and schizophrenia, differs from that of normal subjects [1]. Results of the studies are in accordance with the Crow theory that supposed the causes of mental disorders to be disorders of the mechanisms responsible for the development of lateral asymmetries [2]. It has been confirmed that lateralized training of brain hemispheres can influence behavioural and sensory asymmetries and stimulate one of the hemispheres selectively [3]. These facts strengthen importance to investigate asymmetries in people with mental handicap. It can broaden our understanding of the underlying mechanisms of mental disorders in order to develop better diagnostic and therapeutic methods. Therefore it seems necessary to develop technological instruments fulfilling the following requirements: relative simplicity of use, high effectiveness, interest and motivation for the subjects, ability to scale the effect, ability to record the reactions of the patients.

A device for measuring signal transfer within and between hemispheres, so called Sensopress, has been developed at the Centre for Neuropsychological Research at the University of Trier, Germany [4]. It contains two identical panels allowing both, tactile stimulation and motor response with buttons for the fingers of each hand. The buttons have two functions. They can exert a slight tactile stimulation to a finger, and they can be pressed down by the finger to provide a motor response to the tactile stimulation allowing measuring the response time. This device is controlled by a computer. Among the methods implemented in the software the following were used for the purposes of the present experiment:

Finger tapping test. The subject is asked to repeatedly press a button on the panel as quickly as possible by the index finger of one hand for 10 sec. After a break of 5 sec. the same procedure is performed by the index finger of the other hand. This procedure is repeated 12 times. The device records the mean number of finger taps for each hand as a measure for manual motor speed, and the average time for one finger print is calculated.

Unilateral processing of tactile stimulation. The subject is asked to concentrate attention on the fingers and keep them on the buttons. As soon as a slight tactile stimulation is applied to a finger the subject must respond by pressing the button with the same finger. In this case the tactile stimulation and response are performed on the same side, and this allows measuring the rate of tactile processing within one hemisphere. The response time for the left-sided condition (L/L) indicates the rate of right hemispheric processing whereas response time for the right-sided condition (R/R) indicates the rate of left hemispheric processing.

Bilateral processing of tactile stimulation. As soon as the subject detects tactile stimulation of a finger of one hand he must respond with the same finger of the other hand (R/L and L/R). In this case the tactile stimulation is performed on one side whereas response is performed on the other side, and it means that the stimulus transfers from one brain hemisphere to the other. This procedure allows measuring the rate of interhemisphere transfer by calculating the difference in response time between the bilateral condition and unilateral condition mentioned above.

The experiments were performed on 28 boys with autism spectrum disorders (age: 13.90 ± 0.58 years) and 20 normal boys (age: 12.56 ± 0.65). Only index fingers of both hands were used in these experiments. The experimental procedure started with finger tapping followed by the procedures with unilateral and bilateral processing of tactile stimulation.

Both autistic and normal children responded positively to the test procedure and liked to participate in the experiments. They accepted it as a kind of game.

The results of the finger tapping test have shown that manual motor speed of both right and left index finger did not differ in autistic children compared to the controls, but they differed significantly in the response time to tactile stimulation. Thus L/L time was shorter than R/R in 86.2% of the autistic boys and 33.3% controls ($p<0.01$). The difference L/L – R/R was significant in the autistic subjects, (-25.8 msec. and +5.40 msec. in the children with autism and normal subjects correspondingly, $p<0.05$). R/R time in autistic children was longer than R/R in the controls (299.41 ± 36.50 msec. compared to 205.16 ± 24.36 msec. correspondingly, $p<0.05$).

The results indicate that the left hemisphere processing was slower than that of the right hemisphere in autistic children. Their brains were more asymmetrical, and the left hemisphere conducted the signals slower compared to the normal subjects. It corroborates with the results in the literature and conceptions of other authors [5].

References

1. Burroni, L, Orsi, A, Monti, L, Hayek, Y, Rocchi, R, Vattimo, A.G. (2008). Regional cerebral blood flow in childhood autism: a SPET study with SPM evaluation. *Nucl Med Comun*. **29**, 150-156.
2. Crow, T.J. (1984).A re-evaluation of the viral hypothesis: Is psychosis a result of retroviral integration at a site close to the cerebral dominance gene? *British J Psychiatry* **145**, 243-253.
3. Wittling, W. (1995). Brain asymmetry in the control of autonomic-physiologic activity. In: Davidson, R.J, Hugdahl, K., eds. *Brain asymmetry*. Massachusetts, Cambridge: MIT Press, 305-358.
4. Wittling, W, Wittling, A.R. (2005). Europaische patentschrift EP 1 656 885 B1.
5. Annett, M. (2002). Schizophrenia and autism. In Annett, M. *Handedness and brain asymmetry*. Tailor &Francis Inc., NY, 261-280.

Celeration charts as an assessment tool of learning and executive performance in psychiatric patients

Gwenny Janssen[1], Jos Egger[1,2], Hubert De Mey[2], and Cilia Witteman[2]
[1]*Vincent van Gogh Institute for Psychiatry, Venray, The Netherlands, j.egger@psych.ru.nl*
[2] *Behavioural Science Institute, Radboud University Nijmegen, Nijmegen, The Netherlands*

The term executive functions (EF) has been used to refer to many dimensions of complex human behaviour. It encompasses a broad set of cognitive skills that are responsible for the planning, initiation, sequencing and monitoring of complex *goal-directed* behaviour [1]. For behaviour to be classified as EF, it may not be smooth or automatic but must evoke thinking, problem-solving, and decision making in reaction to *new* or *unusual* situations that conflict systematically with well-established sources of behaviour regulation. Efficient executive performance is thus characterized by the flexibility of an individual to follow, apply, derive or generate rules or strategies in novel situations or, in short, the ability of an individual to *learn* in new or conflict situations.

Learning is a continuous process of acquiring knowledge; in order to form new and more effective rules or strategies, people extract knowledge about problem structures from previous solution attempts they made, but different subjects may learn different rules or strategies [2]. Our aim is to systematically explore (and influence) these individual behavioural differences to offer a broader window on the performance and trainability of executive functioning in clinical and non-clinical groups which, in time, could lead to a better understanding and treatment of executive dysfunctioning.

The Tower of Hanoi (TOH) puzzle has proved to be a suitable task environment in which to study a variety of executive processes [2]. Welsh and Huizinga [3] suggested that the constraints of the TOH are conducive to the spontaneous generation of several problem solving strategies that vary in effectiveness and may explain normal individual differences in performance. Using the TOH as an procedural learning task (administering a single TOH problem repeatedly) will provide information on the issue of *learning*: does repetition lead to substantial improvement of accuracy and speed in performances? How do individuals differ in their ability to learn?

The use of strategy and strategy change was measured by analyzing the acquired verbal reports. This method has considerable utility in the study of complex human behaviour and allows for detailed analysis of human language and cognition from a behavioural perspective [4]. However, competent and adaptive executive performance is not only characterized by the use of efficient strategies, but is also dependent on the fluid combination of accuracy *and* speed of performing. When learners achieve certain frequencies of accurate performance they seem to retain and maintain what they have learned, remain on task or endure for sufficient periods of time to meet real-world acquirements (even in the face of distraction), and apply, adapt or combine what they learned in new situations [5].

To monitor and analyze the growth of learning across time (i.e. the *celeration*), correct and incorrect responses are plotted in a Standard Celeration Chart. This chart is a standard display of frequency as count per time interval (e.g. minute, week, year) and can be used to display change in any human behaviour [6]. Frequency is performance: it tells what happened during one time period, but by itself it tells little about learning. To see whether performance accelerates or decelerates we need to measure it across time [7]. By representing both frequency and celeration in standard graphic and quantitative units, the standard chart clearly differentiates between changes in performance level (frequencies) and changes in learning rates (celerations); higher frequencies of accurate executive performance on a problem solving task predict individual learning ability [5].

References

1. Royall, D. R., E. C. Lauterbach, et al. (2002). Executive control function: a review of its promise and challenges for clinical research. A report from the Committee on Research of the American Neuropsychiatric Association. *Journal of Neuropsychiatry and Clinical Neuroscience*, **14**, 377-405.
2. Simon, H. A. (1975). The functional equivalence of problem solving skills. *Cognitive Psychology*, **7**, 268-288.
3. Welsh, M. C. and M. Huizinga (2005). Tower of Hanoi disk-transfer task: influences of strategy knowledge and learning on performance. *Learning and individual difference*, **15**, 283-298.
4. Cabello, F. and D. O'Hora (2002). Addressing the limitations of protocol analysis in the study of complex human behavior. *International Journal of Psychology and Psychological Therapy*, **2**, 115-130.
5. Binder, C. (1996). Behavioral fluency: evolution of a new paradigm. *The Behavior Analyst,* **19**, 163-197.
6. Pennypacker, H. S., A. Gutierrez, et al. (2003). *Handbook of the Standard Celeration Chart*. Cambridge, Concord.
7. Calkin, A.B. (2005). Precision Teaching: the standard celeration charts. *The Behavior Analyst Today*, **6**, 207-213.

Correlation between mood during daily life and autonomic nervous system activity during sleep

Kohzoh Yoshino and Katsunori Matsuoka
Institute for Human Science and Biomedical Engineering, National Institute of Advanced Industrial Science and Technology (AIST), Osaka, Japan, yoshino-k@aist.go.jp

Sleep is one of the most important activities in our life. Surveys conducted by the National Sleep Foundation reveal that 60 percent of adults report having sleep problems a few nights a week or more [1]. There are many possible reasons for sleeping poorly. Those are mental state (mood), shift work, pain, illness, drugs, and so on. It is known that parasympathetic nervous system activity dominates sympathetic nervous system activity during deep Non-REM sleep [2]. Brosschot et al. showed that daily worry is associated with high heart rate and low heart rate variability (sympathetic dominance) during the subsequent nocturnal sleep period [3]. To our knowledge, there are no other studies that had investigated the relationship between various types of mood other than worry during daily life and the autonomic nervous activity during subsequent sleep. In this study, we examine how eight types of mood during daily life affects the activity balance of autonomic (parasympathetic and sympathetic) nervous system during the subsequent nocturnal sleep.

39 male subjects participated in experiment after provided their informed written consents. The experiment was approved by the Ethical Committee in National Institute of Advanced Industrial Science and Technology. Subjects wore wearable device that measures heart rate variability (RR-interval) and body acceleration signal during daily life including sleep for 72 hours intermittently except when taking bath or shower. Subjects answered to visual analog scale (VAS) questionnaire every hour during waking. The questionnaire scales subjects' eight types of mood level. Those are happiness, tension, fatigue, worry (anxiety), depression, anger, vigor, and confusion. We estimated the activity balance of parasympathetic nervous system and sympathetic nervous system from heart rate variability signal. This was done by calculating HF/(LF+HF), where HF and LF denote the powers of heart rate variability signal in high frequency band (0.2-0.5Hz) and low frequency band (0.04-0.15 Hz), respectively. We determined the timing when subjects had gone to bed (denoted as GBT) by their subjective report and body acceleration signal. We calculated the mean values of each mood measured from six hours before GBT to GBT. We also calculated the mean values of HF/(LF+HF) measured from GBT to three hours after GBT. In order to reduce the inter-individual difference effect, we normalized the values of mood level and HF/(LF+HF) before taking their mean values, and moreover, we took the difference of the mean values in the second night from those in the first (reference) night, before comparing mood level and HF/(LF+HF). The correlation coefficient between depression level from six hours before GBT to GBT and HF/(LF+HF) from GBT to three hours after GBT was -0.59 (see Figure 1 (a)). Moreover, the correlation coefficient between worry (anxiety) level from six hours before GBT to GBT and HF/(LF+HF) from GBT to three hours after GBT was -0.50 (see Figure 1 (b)). These results imply that both depression and worry (anxiety) before sleep shift balance of autonomic nervous system towards sympathetic dominance during sleep.

References

1. National Sleep Foundation. *Sleep in America Poll* Available from http://www.sleepfoundation.org/
2. Trinder, J. (2007) Cardiac activity and sympathovagal balance during sleep. *Sleep Medicine Clinics* **2**, 199-208.
3. Brosschot, J.F., Van Dijk, E., Thayer, J.F. (2007) Daily worry is related to low heart rate variability during waking and the subsequent nocturnal sleep period. *International Journal of Psychophysiology* **63**, 39-47.

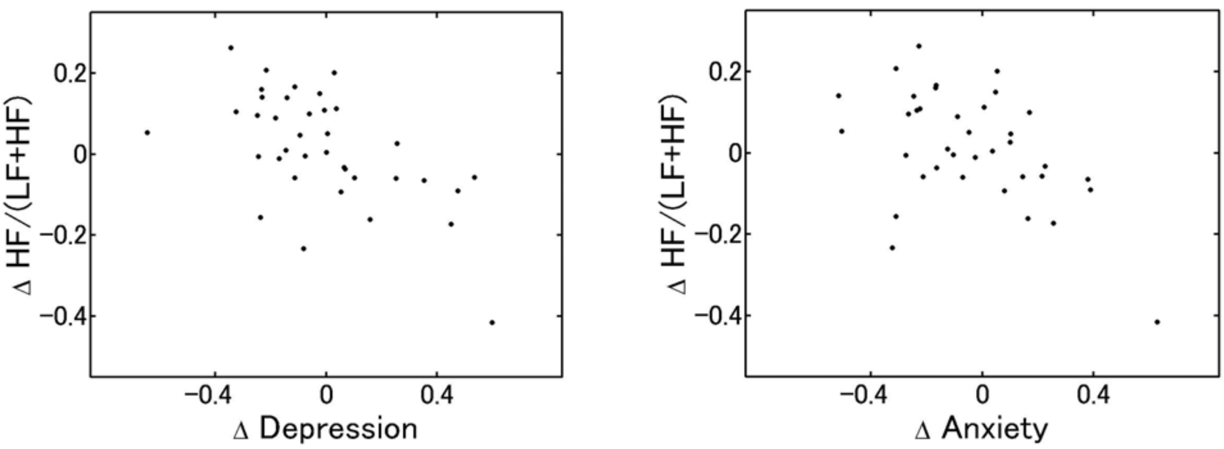


Figure 1. *Correlation between autonomic nerve activity balance index HF/(LF+HF) and mood states, (a) vs. depression, R=-0.59, (b) vs. worry (anxiety) R=-0.50.*

New wireless portable physiology monitor for psychophysiology research

Thijs Schrama
Faculty of Social Sciences, Leiden University, The Netherlands, tschrama@fsw.leidenuniv.nl

Introduction

Several research groups at the Leiden University Faculty of Social Sciences conduct psychophysiological research. In this type of research, physiological measurements are taken from participants during psychological research. Researchers may simultaneously measure the electrical activity of the heart with ECG, Galvanic Skin Response (GSR) measures the electrical impedance of the skin, Breathing rate, and the chest cavity impedance using external sensors with Impedance cardiogram (ICG). These signals represent different ways to measure the physiological state of a person and can be influenced by levels of concentration and arousal.

Currently, off-the-shelf equipment is available to measure physiological responses. However, this equipment has several drawbacks. Most devices do not have online display capabilities. This may result in unnoticed bad signal integrity if electrodes become disconnected. Moreover, some of these devices transmit signals via an analog radio link to the base station, where it is then digitized. This leaves the signal vulnerable to degradation during the analog radio transmission.

Development of an improved Wireless Physiology Monitor

I investigated the possibility to use commercially available personal digital assistants (PDAs) as an alternative platform for such devices. Currently, commercially available PDAs combine processing power, displaying capabilities, and digital wireless connectivity. They provide a excelent platform for taking measurements with on-screen signal visualization and digital wireless data transfer at a low cost.

The system I developed contains an online graphical display and digital transmission of the full signals. This means that researchers can check the signal integrity on the PDA screen and store the entire true waveforms of the signals on both the PDA and the base station for later offline analysis. To facilitate offline analysis, researchers can synchronize experiments with data because the system allows them to place and store markers with the data.

Wireless capabilities

The system makes use of the WiFi network capabilities to send the acquired data to a PC based client application. This digital wireless communication does not introduce any data degradation as apposed to analog data transmission. If the digital wireless connection is lost, the data is logged to memory until the connection is restored.

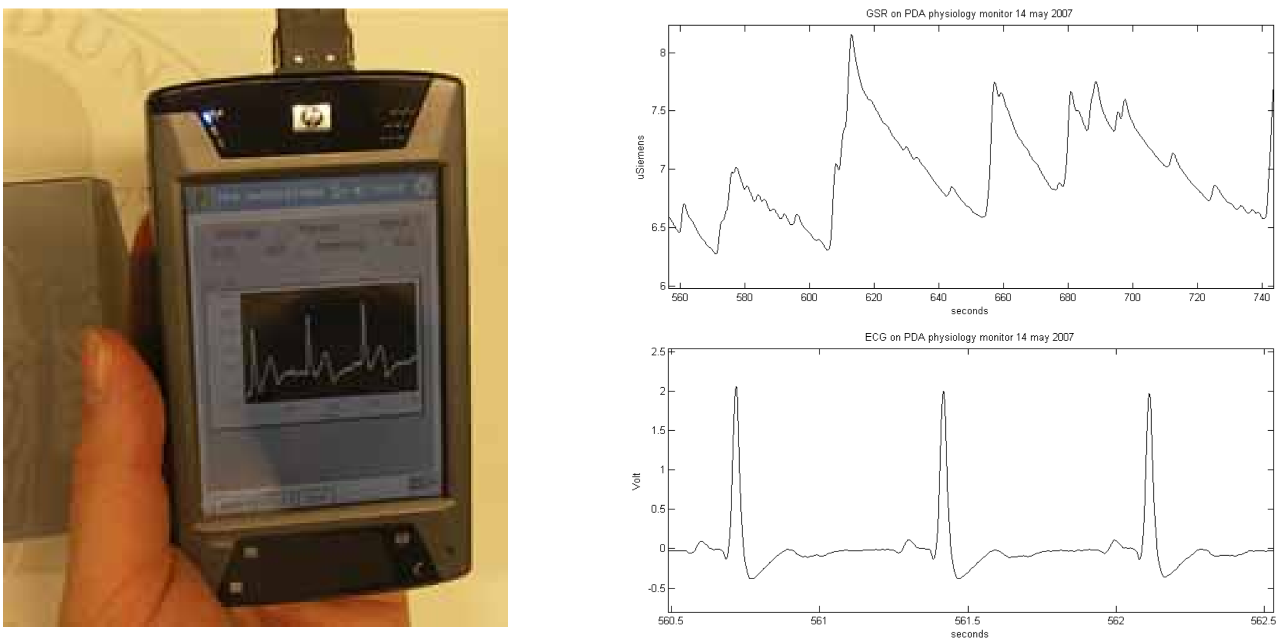


Figure 1. *Left side:Photo of the PDA based Physiology monitor; Right side: Examples of high qaulity data acquired with the PDA based Physiology monitor*

Proceedings of Measuring Behavior 2008 (Maastricht, The Netherlands, August 26-29, 2008)
Eds. A.J. Spink, M.R. Ballintijn, N.D. Bogers, F. Grieco, L.W.S. Loijens, L.P.J.J. Noldus, G. Smit, and P.H. Zimmerman

Conclusion and future work

The small ambulant physiological monitoring system has several advantages over existing equipment. The portable device features:

- Direct online display of the measured ECG, GSR data
- 2 additional analog channels free for custom use
- Onboard logging to file
- Wireless digital transmission of the full data stream to a host computer
- Lossless reconnect capability if the wireless link is lost
- Considerable cost saving over comparable systems, which do not contain online graphical displays or digital telemetry of the signals being measured

After initial tests we plan to use multiple devices in different departments of our university. Our future research may focus more on adding online and offline data analysis tools to the software application.

A Qualitative and Quantitative Analysis of Teacher-Students' Classroom Interaction Communication

Shin-Yin Chien

National Taitun University, Departement of early childhood education, Taitung, Taiwan, linyu8888@nttu.edu.tw

Abstract

This study aimed at investigating learning performance and changing progress of eight experienced teachers in Taipei city public elementary schools, who employed the professional development curriculum of reading strategy instruction. Observed in a 5th grade language and literacy studies class, adapted a mixed method design as well as a concurrent triangulation design. Data collection included classroom observations, video and audio taping, in-depth interview with the participating teachers. The Observer software was used to code and analyze the participants' task.

The results indicated that the teachers who took part in this professional development program did have significant improvement on teachers' pedagogical practice strategies and further on students' reading performance. From the dimension of teachers instruction, the focus of participating teachers shifted from teacher-centered and single instruction with an emphasis on meanings of individual words or phrases to student-partaken in classroom discussion. The findings of the study demonstrated that under the interaction among workshops of in-school reading strategy instruction, teachers' pedagogical practice and teaching beliefs, and students learning performance, the teachers did undergo changes on the process of reading strategy instruction. Students started to participate in communicative discussion in classroom activities, and teachers also shifted from content explanation to meaning construction.

However, the communicative construct during the class still tended to be initiated by teachers and responded by students. Through adjustments of cognitive conflict of interactions with colleagues, teaching demonstration of peer tutors, pedagogical practice improvement and the in-depth interview with the researcher, the participating teachers were motivated to explore their teaching problems, modify and reconstruct their teaching models. The study induced possible learning theories of on-job teachers and modified the pedagogical practice of the professional development via the changing processes of teachers' learning. The findings can facilitate on-job teachers to plan and implement professional development program, which has been proved to be able to meet the needs of teachers' learning characteristics of such programs.

Keywords

Teacher learning, reading strategy instruction, peer tutoring, modeling

Affective valence and activation of the 196 Spanish fragments from the Dasí, Soler y Ruiz (2004) database

C. Dasí, M.J. Soler, J.C. Ruiz, and I. Fuentes
Faculty of Psychology, University of Valencia, Valencia, Spain, carmen.dasi@uv.es

There are four groups of factors that should be considered when studying human memory: subject, encoding, retrieval, and material variables [3]. The last factor is related to several stimulus features that can affect the performance of subjects in a memory task. The correct choice of the experimental material is important because it can affect the validity of the research. One of today's most extensively used implicit perceptual tasks is the word-fragment completion test [8]. The availability of standardized information about fragments can help in the selection of appropriate stimuli. In the Spanish language there is a published normative database of 196 word fragments including some indices like difficulty, priming, familiarity, frequency, etc. [2].

This work aims to extend the available information in this base adding two new indices, "affective valence" and "activation", that can be relevant in affective priming research. The recent dimensional theories of emotion propose that human emotions are based on some basic dimensions. Two of those relevant dimensions are affective valence (pleasantness) and activation (excitation) [4,6]. It can be interesting to have information about these factors to promote the research of affective priming in implicit memory tasks like fragment completion. The control of the affective valence and activation of the stimuli is necessary if a researcher wants to guarantee the validity of the design.

We have obtained both indices for the words of the Dasí, Soler & Ruiz [2] database using a visual scale called SAM [4] ranging from 1 (no affective valence or activation) to 9 (maximum valence or activation). The words were scored in activation and valence by different groups of Psychology students. Each word was assessed by an average of 148 participants. A regression analysis was performed to study the best fitting relation between both dimensions. The quadratic model (boomerang figure) was the best explaining the relation between affective valence and activation, corrected R^2=.26, $F(2, 193) = 34.30$, $SMe = 0.72$, $p < .001$. Regression equation was $x_1' = 10.47 - 2.08x + 0.19x^2$.

We formed two stimuli groups (pleasant-unpleasant) using the median (5.03). With unpleasant stimuli we confirmed the negative correlation between valence and activation (R=-.71; p<.001) obtained by previous researchers [5]; that is, the stimuli most unpleasant produces more activation. However, with pleasant stimuli we obtained a positive correlation (R=.31; p<.01) that differs of other previous results [5,7].

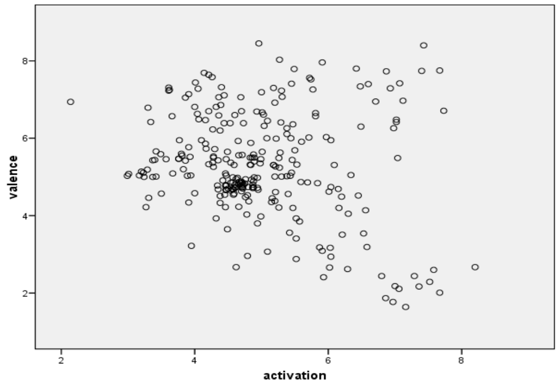


Figure 1. *Distribution of the valence and activation evaluations of the 196 fragments.*

References

1. Bradley, M.M. y Lang, P.J. (1999). *Affective norms for English words (ANEW): Instruction manual and affective ratings*. Gainesville, FL: The Center for Research in Psychophysiology, University of Florida.
2. Dasí, C., Soler, M.J. y Ruiz, J.C. (2004). Normative data on the familiarity and difficulty of 196 Spanish word fragments. *Behavior Research Methods, Instruments, & Computers,* **36**, 559-563.
3. Jenkins, J.J. (1979). Four points to remember: A tetrahedral model of memory experiments. In L.S. Cermak & F.I.M. Craik (Eds.), *Levels of processing in human memory* (pp. 429-446). Hillsdale, NJ: Erlbaum.
4. Lang, P.J. (1980). Behavioral treatment and bio-behavioral assessment: Computer applications. En J.B. Sidowski, J.H. Johnson y T.A. Williams (Eds.), *Technology in mental health and delivery systems* (pp. 119-137). Norwood, NJ: Ablex.
5. Moltó, J., Montañés, S., Poy, R., Segarra, P., Pastor, M.C., Tormo, M.P., Ramírez, I., Hernández, M.A., Sánchez, M., Fernández, M.C. y Vila, J. (1999). Un nuevo método para el estudio experimental de las emociones: el International Affective Picture System (IAPS). Adaptación española. *Revista de Psicología General y Aplicada,* **52**, 55-87.
6. Osgood, C., Suci, G. y Tannenbaum, P. (1957). *The measurement of meaning*. Urbana, IL: University of Illinois.
7. Redondo, J., Fraga, I., Comesaña, M. y Perea, M. (2005). Estudio normativo del valor afectivo de 478 palabras españolas. *Psicológica,* **26***,* 317-326.
8. Toth, J.P. (2000). Nonconscious forms of human memory. In E. Tulving & F.I.M. Craik (Eds.), The Oxford handbook of memory. Oxford, UK: Oxford Univerisity Press. Pp. 245-251.

1

The “ANS-Explorer”: A newly developed instrument for measuring autonomic regulation ability by analyzing heart rate variability

R. A. Wittling and E. Schweiger
Center for Neuropsychological Research, University of Trier, Trier, Germany, arne-wittling@znf.uni-trier.de

The autonomic nervous system (ANS) with its branches sympathicus and parasympathicus provides the organism with a regulatory system which is able to flexibly adapt to the demands of the body’s organs. While the sympathetic system usually activates bodily processes to prepare them for appropriate dealing with stress situations the parasympathetic system usually has the antagonistic function in order to relax and restore bodily processes and to regain functional balance of the organs and to maintain homeostasis. In general, it is of great importance that these both branches act flexibly to increase and decrease nervous activity. Disturbances of the ANS such as high levels of sympathetic activity of low levels of parasympathetic activity are directly involved in a multiplicity of somatic and psychosomatic illnesses especially if the cardiac and circulatory system is affected. While sympathetic innervation for example accelerates the beat frequency, parasympathetic innervation slows it down.

Concerning cardiac functions the procedure which is most suitable at present to meet the requirements of measuring sympathetic and parasympathetic activity is the analysis of heart rate variability (HRV). The most important parameters of the HRV are derived from time analysis methods at the one hand and from frequency or spectral analytic procedures (power spectrum analysis) at the other hand. Among time related parameters “SDRR” (standard deviation of all RR times during the measured period) and “RMSSD” (differences between the RR intervals of successive heart beats) are mostly considered. Concerning frequency related parameters “high frequency power” ($HFms^2$; including faster fluctuations in heart rate frequency between 0.15 – 0.40 Hz), “low frequency power” ($LFms^2$; including slower fluctuations between 0.04 – 0.15 Hz), “total power” ($TPms^2$; addition of the power of different frequency bands), “$LFms^2/HFms^2$ ratio” (measures the relative degree of the sympathetic versus parasympathetic activation; sympathovagal balance) are of most importance. Especially $HFms^2$ is considered as a measure of sympathetic activity while $LFms^2$ is considered as a measure of parasympathetic activity.

At the Center for Neuropsychological Research (University of Trier) a special device was developed which fulfils all requirements for measuring HRV. The “ANS-Explorer” consists of three networked modules: ANS-Tracer, ANS-Explorer und ANS-Trend.

The ANS-Tracer module is a patient administration system which receives electrocardiographic data from a medical 3-channel ECG amplifier via a secured bluetooth radio connection and provides these data for further processing in chronological order using all the necessary patient information. The ANS-Explorer as a central module prepares the imported RR intervals and calculates the different time and frequency related parameters of HRV. In order to calculate these parameters and to arrive at a valid conclusion it is necessary to make various mathematical calculations of the measured output data prior to being fed for instance for spectral analysis. The data pre-calculated in this way can be displayed in various ways, for instance some forms of representation are more suitable for practical diagnostic purposes, others for research purposes. Finally all the individual data can also be shown numerically and graphically as deviations from a population norm in order to clarify the individual’s position in a reference group. This method of representation makes it easy to assess the measured values of an individual with respect to age, gender, patient or training group and his ranked position in his comparison group.

Up till now some studies demonstrated the possibility to predict the risk of disease in healthy people on the bases of HRV data such as the mortality risk after myocardial infarction, the occurrence of malignant arrhythmias, sudden cardiac death, and diabetes and autonomic neuropathy. Studies of our research group dealt with, for example, influences of personality traits in normal healthy subjects and in psychosomatic patients on HRV variations. Healthy subjects who are characterized by a flexible HRV show a much more positive prevailing mood, are more socially oriented, less withdrawn, more empathic, and tend to express both positive and negative feelings. In a study on 50 chronic psychosomatic patients compared to healthy subjects it could be shown that psychosomatic patients in general show an impairment in the regulation of ANS activity. Correlated with this impairment are an increased stress reactivity, an increased number of physical complaints, and a generally raised level of psychiatric symptoms. Furthermore, it is important to note that a reduced ability to regulate ANS activity goes along with a higher level of psychological and physiological complaints.

In times of growing stress und stress-related strain we have furthermore developed an additional and innovative version of the ANS-Explorer based on a mobile phone which allows at any time to get information about the personal body’s stress level. This knowledge enables an individual to change stress-related behavior to reduce the danger to suffer from stress-induced diseases, for example sudden cardiac death . This system determines fast and reliably the individual’s autonomic level. The device consists of a common mobile phone with an integrated or blueblooth-connected ECG which measures heart activity by finger electrodes. Analysis of HRV and the presentation of the individual autonomic regulation state is presented immediately after recording. It is possible for the user to have a look at changes of his stress-related autonomic processes any time and directly at the mobile phone or via an internetportal.

Taken together, HRV measurement seems to be an innovative and appropriate tool for diagnosis and prognosis of health and disease. Concerning these questions the “ANS-Explorer” is a tool which fulfils all necessary methodological and statistical requirements to answer these questions.

References

1. W. Wittling, E. Schweiger & R. A. Wittling (2008). Diagnostik der Herzratenvariabilität: Einblicke in die autonom-nervöse Regulation von Stressverarbeitung, Befindlichkeit, Verhalten und Gesundheit. Research Report, Center for Neuropsychological Research, University of Trier, Germany.

The Analysis of User Behavior observation at open café in downtown: Focused on the open space in downtown

Kyung Ran Choi[1] and Eun Jung Lee[2]

Graduate School of Techno Design, Kookmin University, Seoul, South Korea,

[1]ran@kookmin.ac.kr, [2]freedom44s@naver.com

The open space in downtown provides some rest places and cultural living space to the citizen. It's a part which is considered importantly for users that a cafe is a place where the people gathers. Like this, a cafe has become the rest space of the moderns as part of a new trend that differ from the rest like a park or a square. So I'd like to find out what kind of cultural characteristics, based on the cultural assumptions of the society, is the primary reason that people spend their time in cafes. In short, I'd like to examine aspects of cafes as resting places where people search for repose in the highly urban environment of modern society. Accordingly, this research defines a resting place as an artificial space where people can conduct their internal lives during free time.

First of all, I will search documents and phases associated with relaxation in several cafes or open spaces in the form of pictures and memos; then I'll film one specific place in order to find out people's satisfaction and dissatisfaction from a social and psychological viewpoint. When filming, I'll choose a cafe with some open space and then set up a camcorder where the movement of people is easily seen. I'm going to analyze the aspects and activities of cafes and then provide a new design method based on the results.

With matter found from preparatory research and filming, I am going to show the relationship between a cafe as a resting place and people’s accessibility and gratification. Based on this material, I will analyze detailed aspects by using the Observer program of the Noldus Company. I will offer explanations for the facts that emerge in the course of the filming and make an analysis on the new pattern. This analysis can be a new basis for providing a way of designing.

Based on the results from the film, I will analyze the aspects of a cafés a relaxation area. Then, I am going to lay out problems of design by describing the causes. By observing people using a bench or a chair not only for sitting but also for other activities such as drinking a cup of tea, waiting, or talking, I will see if that bench or chair has a suitable design for resting. The researching method through these analyses suggests that it can play a role in recommending new design methods by using the Observer program, which provides proper analytical tool. Therefore, this study is going to be a guidebook on the convenience of people using a resting place.

Animals' behavioral data analysis using fractal dimension method

Viktor Veliks[1,2] and Alexander Grakovski[2]

[1]Department of Human and animals physiology, Faculty of Biology, University of Latvia, Riga, Latvia, viktor.veliks@gmail.com

[2]Department of Electronic and Telecommunication, Transport and Telecommunication Institute, Riga, Latvia

During statistical analysis of behavioral data it is important to look on behavior as a complex process during observation period. Usually analysis of aggregated observation data (typically by mean or sum values) is conducted using standard statistical procedures and as a result variation of the behavior parameters during experiment is lost, another specific of such way of analysis is that measured behavior parameters are looking "unrelated".

The aim of this study was to develop advanced method of animals' behavior data representation which combined different behavior parameters and time-depending variation of these parameters.

Experimental data from animals' observation in round open field were used for the analysis. Animals' behavior was recorded using SMART software (v 1.2, Barcelona, Spain). The following behavior parameters data were used for calculations – distance, presence in zones and different animal's positions. Values of these parameters were calculated for each second. Data preparation for zones and positions was done using specific coding procedure. Presence in a zone and directions of zone changing were coded by positive numbers.

For activities in zones, a graph of possible animals' movements was used. Encoding, longer distance covered by the animal from one zone to another in a round open field is given higher value. In the graph the first number represents movement in the round open field zones, see figure 1. For example, 131 means that animal stayed in zone 1 then moved in to the zone 3 (center of open field) and returned back in to zone 1. Second number shows value that encodes this movement. Animal's positions were coded with negative numbers for scratching and grooming parameters and with positive numbers for vertical activities, holes observations and walking. Behavior position was recoded using the following values: any vertical position 30 points, holes investigation 15, walking 0 points, scratching -15 and grooming -30 points.

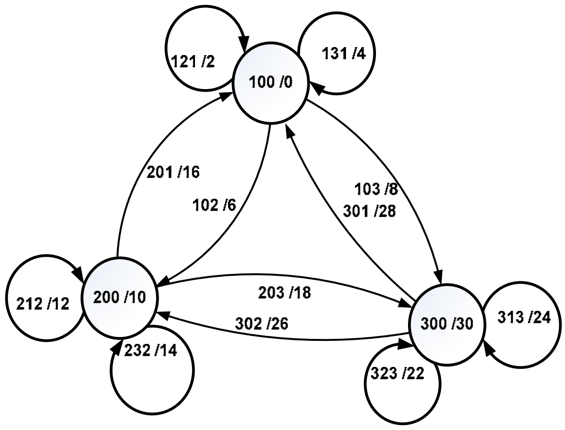


***Figure 1**. Graph of animal movements in the open field zone; first number movement in the zones, second number coding value for this movement.*

After that a 3-D array was created and thereafter used for manipulations. Values of movements in zones and behavior position were divided by mean value of walking distance, as result all scales that represents animal behavior are equal. The special application in Matlab was developed for calculation of fractal dimension using correlation dimension from Grassberger and Procaccia [1].

In the first step of our algorithm 3-D plot (presence in zones, animal positions and distance covered) of observation data was calculated, this gave possibility to conduct visual data analysis. In figures 2 and 3 there is 3-D plots showing two animals' summary of behavior activities comparing to their movement track recorded during the experiment. Each point in 3-D plot represents behavior activity by values of animal positions, zones and distance covered during interval of one second. Figure 2 displays an animal behavior data with a low movement activity. A figure (2.b.) of relatively simple animal behavior phase with one pike on the zone axes can be seen that characterizes the animal entrance in the second zone (2.a.). Also during experiment this animal was less active in

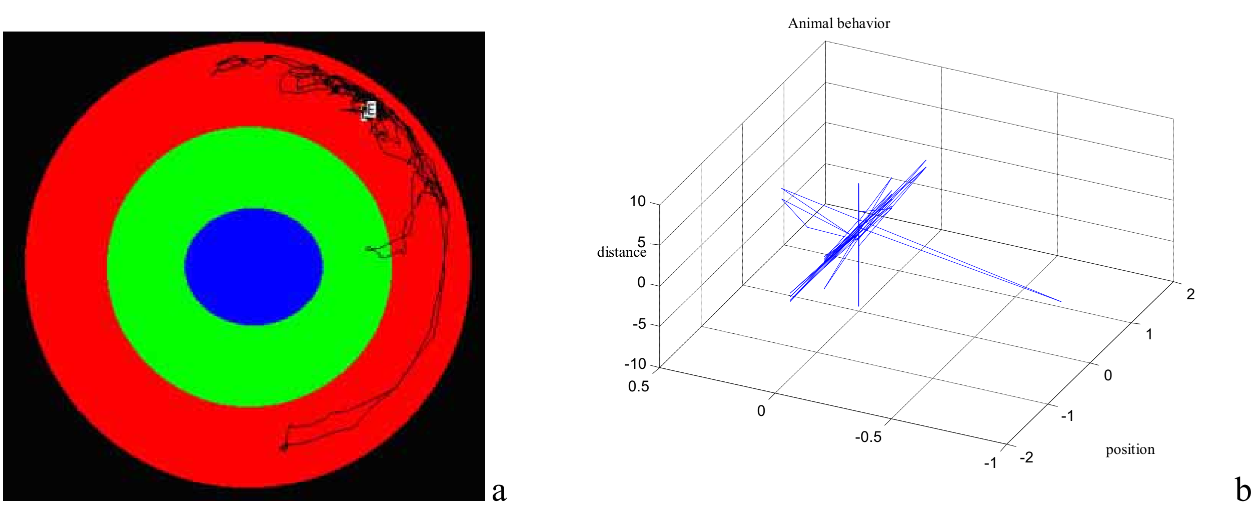


***Figure 2**. Animal with low behavior activity: a. walking path in open field zones, b. 3-D plot of animal behavior.*

Proceedings of Measuring Behavior 2008 (Maastricht, The Netherlands, August 26-29, 2008)
Eds. A.J. Spink, M.R. Ballintijn, N.D. Bogers, F. Grieco, L.W.S. Loijens, L.P.J.J. Noldus, G. Smit, and P.H. Zimmerman

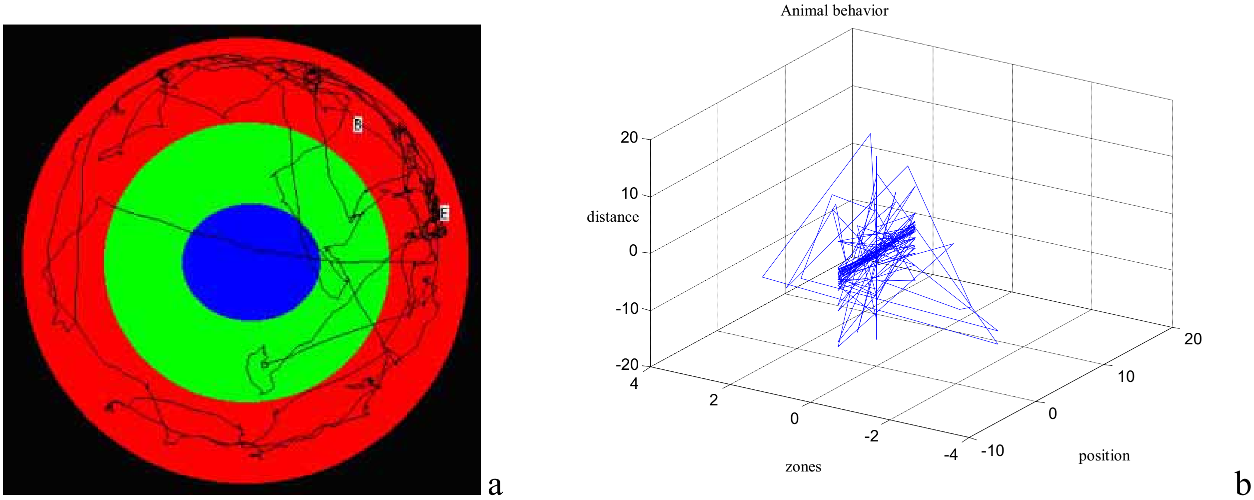


Figure 3. *Animal with high behavior activity: a. walking path in open field zones, b. 3-D plot of animal behavior.*

parameters of position and covered distance.

Figure 3 displays an animal behavior with a very high movement activity. Figure (3.b.) of the behavior phase is already considerably different, the animal actively investigates the territory and even crosses the center of the open field (3.a.) and its behavior is much more varied.

In the second step Euclidean points from 3-D array were calculated. These data used for calculation of value of "correlation dimension". We assume that "correlation dimension" represents complexity of animal behavior. Mathematical formula of the algorithm "correlation dimension" – *Dcor* is the following:

$$Dcor = \lim_{r \to 0} \frac{\ln C(r)}{\ln r}\ \text{, where } C(r) = \frac{\|\, \mathrm{Xi} - \mathrm{Xj} \,\| <= \mathrm{r}}{n^{2}}$$

where, r – length of square side,

||Xi-Xj||<= r – distance between points which are lower or equal to r,

n – number of points in the experimental data,

Dcor – correlation dimension.

Besides, the length of line segment used for calculations – r was taken from the data of animal movements: from the minimal value to the maximal. This method allows using biologically significant step sizes. For fifty r values logarithms were generated, starting with logarithm of minimal to logarithm of maximal value of the walked distance, in order to gain maximally even data representation on the graphic scale.

Finally, fractal dimension was calculated [2]. After finding of $\ln C(r)$ and $\ln r$ values correlation dimension *Dcor* was calculated using regression analysis where *Dcor* is a coefficient in regression equation. We observed that animals with low behavior activity had lower correlation (fractal) dimension values. For example, less active animal fractal dimension *Dcor* value is 0,33264 and for more active animal *Dcor* is 0,59863. That means that animal with low behavior activity (short covered distance, zones and taken positions) has also the simplest behavior patterns.

The results of obtained fractal dimension data may be investigated by standard statistical observation method, for example, by visual method, correlation or clusterization.

This method allows evaluating animal's behavior overall. Animal's behavior data are visualized using 3-D plots, but from the other side "correlation dimension" – *Dcor* gives to behavior numerical value. Another plus of this method is that we can see dynamics of animal's behavior.

References

1. P. Grassberger and I. Procaccia (1983). Characterization of Strange Attractors. *Physical Review Letters,* ***50(5)****,* 346-349.
2. Mandelbrot, B. B., *The Fractal Geometry of Nature*, New York: Freeman, 1982.

Advantages and limits of a video multitracking system for quantification of individual behavior in a large fish shoal

J. Delcourt[1], C. Becco[2], N. Vandewalle[2], and P. Poncin[1]

[1]*Behavioral Biology Unit: Ethology and Animal Psychology, Department of environmental Sciences & Management, Faculty of Sciences, University of Liège, Belgium, Johann.Delcourt@ulg.ac.be*

[2]*Group for Research and Application in Statistical Physics (G.R.A.S.P.), Department of Physics, Faculty of Sciences, University of Liège, Belgium*

Introduction

Video multitracking systems able to track animal groups have recently been developed but are not yet in common use. To study fish shoals, the group sizes exceed rarely some individuals. This is insufficient to study schools of real size, which often contain hundreds of fish. Moreover, when the trajectories of two fishes cross each other and their images coincide (termed occlusion), these programs are usually unable to attribute the correct identity to each fish after they uncross, or separate. Others individual multitracking systems able to track a large number of targets have been produced. Unfortunately, a similar system does not yet exist for fish tracking. Moreover, the behavior of fish is more complex because the apparent image of the fish is highly variable and the direction and speed of fish displacements can change very quickly in terms of time and magnitude.

Delcourt *et al.*[1] have developed a 2D multitracking system to quantify the schooling behaviors in fish using a laboratory approach. A dynamic predictive model is based on the previous measures of velocity vectors [2]. The system's ability to correctly track a large number of individuals (up to 100) in the same arena is introduced.

Materials and methods

We worked with images converted to JPEG format and eliminated the background noise with a reference image of the arena with no fish. Then a range of grey-scale characteristics of the image of the fish was detected. The detected pixels were converted to a binary format that was usable by our tracking program.

We estimated the capacity of our system to assign the correct identification to each fish compared with the human eye. We define the Recognition ratio of individual fish = $A*T_O + B*T_S + C*T_N$.

(T_O = number of identity assignments in the context of occlusion/total number of identity assignments; T_S = number of identity assignments when separation/total number of identity assignments; T_N = number of identity assignments in other cases /total number of identity assignments; A = successful identification ratio when there is occlusion; B = successful identification ratio when there is separation; C = successful identification ratio in other cases).

Results

The ratio of successful identifications varies in relation to the identification events (see table 1). The greatest difference appears between the cases of occlusions and no occlusions. When there are no occlusions, the program is very efficient (close to 100%). With a good image rate, if there is no contact between the images of two fishes, there is no problem identifying each fish. When occlusions occurred, two situations were distinguished: during the occlusion and the subsequent separation of the images of the two fish (end of occlusion). During the occlusion, we considered that the spot (the image of the two superimposed fish) has a double identity from both fish, making the identification nearly always correct. In the separation situation, the ratio of successful identifications ranged from somewhat good (83%) to poor (close to 50%). This was the main cause of identification errors. Two situations could occur.

In the first situation, the program did not detect the separation. A fish retained the double identity and the second fish had no identification number. The second was an orphan if the program searched for a constant individual number. When the program provided for the appearance of a new individual during the tracking, a new identification number was assigned to the orphan fish. In this case, the identification error was easy to detect. Moreover double identity (occlusion) is a short periodic state. If a fish retains a double identity over more than 0.5s (see figure 2), double identity can be considered uncertain. Knowing this characteristic makes it possible to automatically detect misidentification when the program fails to detect a separation.

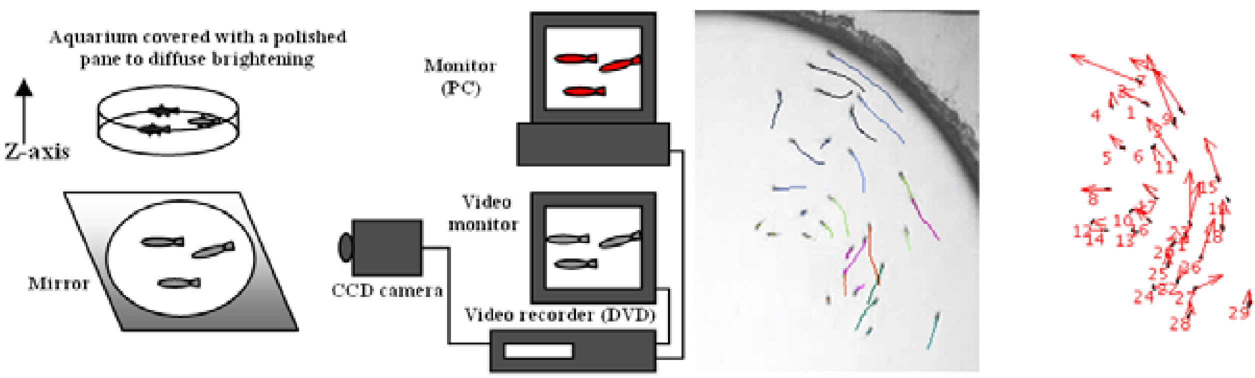


Figure 1. *Left: Diagram of the multitracking setup. A CCD camera filmed a mirror reflecting the bottom of the circular tank. Right: example of tracks (detail of the arena, period of 1 sec) and the same image with individual assignments and velocity speed.*

Table 1. *Ratio of successful identifications in relation to the number of fish (n= number of identification events).*

Nb of fish	Ratio of successful identifications			
	Occlusion	Separation	Others	Total
10	100 (*n* =156)	75 (*n* =21)	99.97 (*n* =9822)	99.92 (*n*=10000)
20	100 (*n* =277)	73.3 *n* =40)	99.99 (*n*=19683)	99.93 (*n*=20000)
30	100 (*n* =259)	71.4 (*n* =37)	99.96 (*n*=29704)	99.93 (*n*=30000)
40	100 (*n* =203)	63.6 (*n* =29)	99.97 (*n*=39767)	99.94 (*n*=40000)
50	100 (*n* =1035)	68.4 (*n*=149)	99.97 (*n*=48816)	99.88 (*n*=50000)
60	100 (*n* =554)	50 (*n* =40)	99.22 (*n*=59406)	99.20 (*n*=60000)
70	100 (*n* =1219)	72.7 (*n* =176)	99.98 (*n*=68605)	99.92 (*n*=70000)
80	100 (*n* =998)	55.5 (*n* =144)	99.78 (*n*=78858)	99.71 (*n*=80000)
90	100 (*n* =1214)	44 (*n* =220)	99.89 (*n*=88566)	99.75 (*n*=90000)
100	100 (*n* =993)	62.5 (*n* =180)	99.88 (*n*=98827)	98.83 (*n*=100000)

Second, identities could be inversed: after separation, each fish retained the identity of the other individual. In this case, it was very difficult to automatically identify the error.

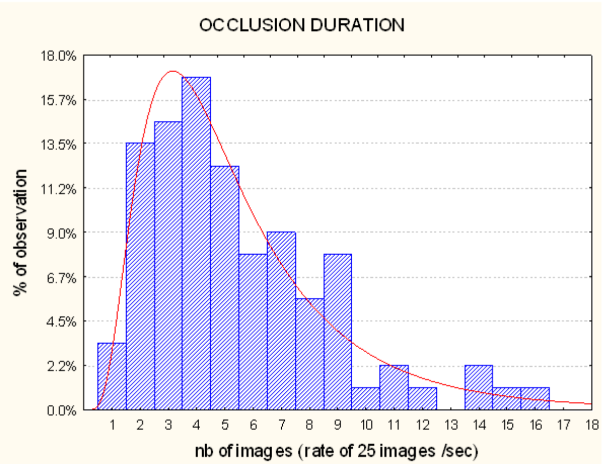


Figure 2. *Occurrence of occlusion duration. The rate of images is 25 images per second. The fit corresponds to a lognormal curve (μ=1.52, σ = 0.61). (n=200).*

Discussion

In statistical tracking, in which the identity of a fish is important for only a short time (e.g., to measure the individual's speed), but not in the final analysis, the program is very efficient, with more than 99% of the individuals successfully identified. On the other hand, for longer duration rigorous individual tracking, a single assignment error has a dramatic effect on the study. Visually verifying the trajectories after crossing events is feasible when fish density is low, but becomes significantly less so as the numbers of fish increase. The computer can help the experimenter by indicating the occurrence and position of a crossing event. However the separation rate is generally less than 0.1% of the total number of attributions and the visual expertise is not a significant work.

Consequently, in both statistical and rigorous individual tracking, this system allows the experimenter to gain more time by measuring the individual position automatically. It can also analyze the structural and dynamic properties of an animal group with a very large sample, with precision and sampling that are impossible to obtain with manual measures.

References

1. Delcourt, J., Becco, Ch., Vandewalle, N. & Poncin, P. (2006). Comparing the *EthoVision*® *2.3* system and a new computerized multi-tracking prototype system to measure the swimming behavior in fry fish. *Behavior Research Methods* **38(4)**: 704-710.
2. Delcourt, J., Becco, Ch., Vandewalle, N. & Poncin, P. (2005). A new automatic video multitracking system able to follow a large school. *Proceedings of Measuring Behavior 2005, 5th International Conference on Methods and Techniques in Behavioural Research* (*Wageningen, 30 August- 2 September 2005)*, pp.405-406.

A new treatment regime for sodium azide to evoke experimental Alzheimer's disease for pharmacological screening

K. Megyeri, M. Albert, H. Kompagne, L. G. Hársing jr., I. Gacsályi, and Gy. Lévay
Department of Pharmacology, Division of Preclinical Research, EGIS Pharmaceuticals Plc, Budapest, Hungary, megyeri.katalin@egis.hu

Alzheimer's disease (AD) is characterized by a progressive decline of cognitive functions. Comparative studies identified that there are no detectable signs of drug-induced changes in cognitive functions in healthy animals brain versus diseased brain, therefore, to see decreased behavioral manner of AD suffered animals needs to mimic all specific features of the disease [1]. More evidence suggests the early role of reactive oxygene species in AD major source of which is the mitochondria [2]. The mitochondrial poisoning hypothesis of AD underlying is based on the fact that in human post-mortem AD brains the complex IV activity declines [3]. Inhibition of this complex could be evoked by chronic subcutaneous sodium azide (NaN_3) treatment via implanted osmotic minipumps in animals [4]. For screening, however, minipumps are not ideal tools due to high cost, one-time usability and long treatment time. We developed a new method to produce AD–like dementia by selective inhibition of cytochrome C complex using intraperitoneally (ip.) injected NaN_3 in various doses (10–15 mg/kg/day) in rats. We explored the dose-effect relationships of the ip. applied mitochondrial poison and established a complex test system to study alterations of the cognitive functions caused by NaN_3 treatment. Furthermore, we found the optimal dose and treatment regime of NaN_3 to evoke histopathological changes in treated animals.

The mechanism of spatial learning behavior in mammals is critically dependent on the proper function of the hippocampus. Based on these data we concluded that the hippocampal CA1– and CA3–regions are crucial for acquisition and memory consolidation therefore, the Morris water maze test was selected to examine the changes in hippocampal-dependent cognitive functions [5]. We measured learning function after 5 days ip. treatment using the TSE VideoMot2 system. This is a system for the camera-monitored, computer-supported, automated observation and analysis of activity and locomotion of animals. It captures the movement of animals in the pool, collects, analyzes the data and draws the swimming path.

Methods

We used male Sprague-Dawley rats (400–430 g) from our own breeding stock. Animals received extruded complete diet and water, ad libitum. Our Local Committee based on the 86/609/EEC directive approved the protocol. During acquisition trials animals placed into the water were only able to escape by finding a platform within 120 s, which was hidden 1 cm under the water in center of one of the quadrants of the maze (target). Each rat received 4 trials/day for 3 days. The trial ended automatically when the target was reached and the elapsed time was called "escape latency". After 14 days NaN_3 treatment we studied the memory functions. We removed the platform from the pool and animals had only one trial (120 s). We registered three parameters: the time spent in the original target quadrant of the pool (1st), the number of crossings through the original position of the platform (2nd), and the latency of the first crossing of the platform original location (3rd). When 12.5 mg/kg/day NaN_3 dose was reduced after 5 days to 10 mg/kg/day, considerable loss of memory functions was revealed on this day, specifically a significant decrease in the 1st and 2nd and increase in the 3rd measured memory parameter. One hour later we tested the re–learning capability of the animals by putting back the platform into a new quadrant of the pool. Animals had re–learning trials from three directions. In this experiment only animals treated with 12.5 mg/kg/day NaN_3 than reduced to 10 mg/kg/day had significantly longer escape latency. To investigate NaN_3-induced interactions we measured the animals' spontaneous locomotor activity (SMA), and the weight of their body and adrenal glands.

Results and Discuission

According to swimming velocity results the observed effects developed without the loss of movement ability. 24 hours after the last NaN_3 treatment SMA, body and adrenal glands weights of animals in any of the NaN_3-treated group did not differ from the control, indicating the lack of any non-specific effect (e.g. stress). Measurement of corticosteron-regulated immunological markers (IL-8, IL-10) is in progress. Detailed histopathology of the different regions of brain was performed at the termination of the study. Neuronal degeneration and necrosis were seen in the cortical and hippocampal areas in the treated rats. Pathological changes in the ultrastructure of mitochondria and glial cells were detected in NaN_3–treated animals suggesting some degree of disruption of blood–brain barrier. Significant loss of neurons was detected in the hippocampal CA–1 and CA–3 regions. Ki67 immunochemistry to evaluate neurogenesis demonstrated decreased rate of cellular proliferation in the dentate gyrus of the treated rats.

These results show that treatment regime of 12.5 mg/kg/day ip. NaN_3 for 5 days with a subsequent reduction to 10 mg/kg/day for the following 9 days is a suitable method to produce dementia in rats.

We confirmed that with the ip. injection method of NaN_3 for 15 days produces comparable level of dementia caused by 31 days infusion of NaN_3 using implanted osmotic minipumps in rats. In summary, our developed new treatment regime seems to be an improved method for pharmacological screening of neuroprotective compounds, where low cost, reproducibility and short experiment time are required.

References.

1. Tsuji, T., Shimohama, S. (2001). Analysis of the proteomic profiling of brain tissue in Alzheimer's disease. *Dis Markers* **17**, 247-257.
2. Smith, T. S., Bennett, Jr. J.P. (1997). Mitochondrial toxins in models of neurodegenerative diseases. I. *Brain Res* **765**, 183-188.
3. Parker, Jr. W.D., Filley, C.M., Parks, J.K. (1990). Cytochrome oxidase deficiency in Alzheimer's disease. *Neurology* **40**, 1302-1303.
4. Bennett, M.C., Mlady, G.W., Kwon, Y.H., Rose, G.M. (1996). Chronic in vivo sodium azide infusion induces selective and stable inhibition of cytochrome c oxidase. *J Neurochem* **66**, 2606-2611.

5. Bennett, M.C., Rose, G.M. (1992). Chronic sodium azide treatment impairs learning of the Morris water maze task. *Behav Neural Biol* **58**, 72-75.

Changes in feeding behavior after high gravity loading in orchidectomized and ovariectomized rats

M. Kimoto[1], Y. Kumei[2], J L. Zeredo[3], and K. Toda[1,3]

[1]*Physiological Laboratory, Japan Women's University, kimoto@fc.jwu.ac.jp,* [2]*Section of Biochemistry, Tokyo Medical and Dental University,* [3]*Section of Integrative Sensory Physiology, Nagasaki University, Japan*

We previously reported that feeding behavior is modulated by stressful high gravity loading [1-3]. In these cases, gender difference is observed, suggesting that sex hormones are highly concerned with changes in feeding behavior after gravity stimulation. In the present study, we investigated the role of sex hormones on the changes in feeding behavior under the condition of high gravity loading. Wistar rats received orchidectomy or ovariectomy at the 53rd day after birth. Rats were exposed to 2G on a centrifuge for 10 min everyday during the period from the 63rd through 92nd day. In male rats, body weight gain was suppressed immediately after orchidectomy. The body weight gain was deteriorated by 2G loading in the orchidectomized rats worse than in the non-orchidectomized rats. The deterioration of body weight gain was initiated in the orchidectomized rats on the 14th day, whereas it was initiated in the non-orchidectomized rats on the 18th day following 2G loading. On the 92nd postnatal day when 2G loading was discontinued, the serum corticosterone levels in either orchidectomized or non-orchidectomized rats was significantly higher than the respective non-loading controls. In female rats, body weight gain was significantly potentiated on and after the 14th day following ovariectomy. However, body weight gain was deteriorated on and after the 13th day following 2G loading in ovariectomized rats. The serum corticosterone level was not influenced by gravity loading both in the ovariectomized and non-ovariectomized rats. These results show that 1) sex hormone is involved in body weight control:increasing effects in male and decreasing effects in female, 2) the resistance against high gravity stress is more potent in female than in male,3) the female hormone is critical for anti-stress behavior in the body weight control system.

This study was supported by a Grant-in-Aid for Scientific Research from the Japan Society for the Promotion of Science to M.Kimoto (No.18500630, 2006-2007). We thank K.Iwasaki, H. Okubo, A. Sakabe, Y. Sakitani,T. Nakajima, C.Yanagisawa and M. Yamashita for analyzing data. The experimental procedures were in agreement with the Animal Care Standards of Nagasaki University, and had the approval of its Animal Welfare Committee (no.0507050445).

References:

1. Kimoto M.,Amano R.,Ishikawa R.,Ooe Y.,Shimura M Niinuma K.,Hirakawa H.,Kumei Y.,Toda K.(2007) Effects of stress applied at the Late stage of pregnancy on the feeding behavior in the rat. *Cell Stress & Chaperons*, **12**, 440-441.
2. Kimoto M.,Takase S.,Toda K.,Kumei Y:(2007) Chemical sensation modulated by high gravity Stress at different growth stage in rats. *Gravity Space Biol*., **21,** 34-34.
3. Kimoto M, Takase S, Shomura E, A. Suzuki A, Zeredo JL, Sasaki-Zeredo K.M, Kumei Y, Toda K.(2005) High gravity modifies feed efficiencies and chemical sensation in rats. *Autonomic Neurosci.,* **110**, 118-118.

Optimization and robustness of the inhibitory avoidance test for pharmacological screening on long-term memory

E. Detrait, E. Hanon, B. Dardenne, and Y. Lamberty
UCB, s.a. Centre for CNS Innovation, Psychopharmacology dpt, Chemin du Foriest, B-1420 Braine-l'Alleud, Belgium,
eric.detrait@ucb-group.com

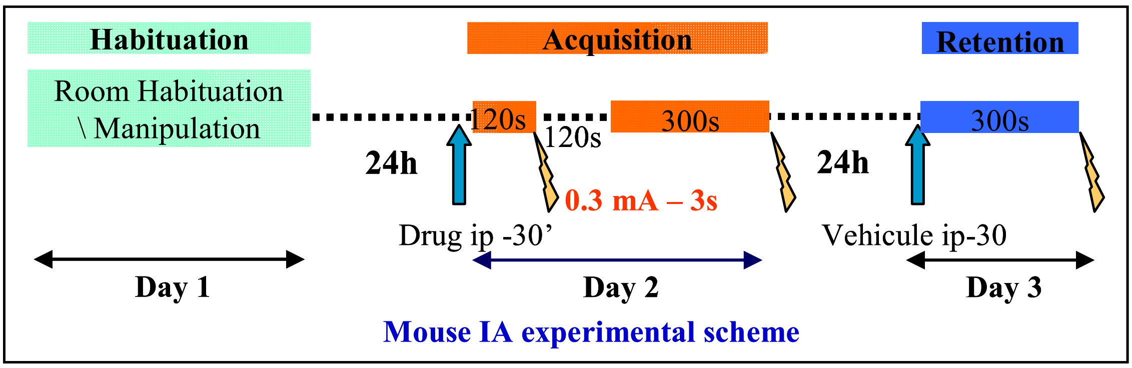
One of the challenges in pharmacological screening is to maximize the throughput of tests in order to evaluate the potential efficacy of a large number of compounds in a minimal amount of time. The inhibitory avoidance test, also called passive avoidance, has long been used as a screening test to evaluate drug effects on the memory in rodents [1,2]. The test is based on the natural photophobia of mice or rats, and evaluates the long-term memory of animals. The apparatus (Ugo Basile, see picture below) consists of a box divided into two compartments of equal sizes (18x10x16 cm) and equipped with a grid floor. One compartment is made of white panels, and illuminated with a lamp placed on the top of the chamber (≈ 350 lux). The other compartment is made of black panels (≈4 lux). The two compartments are separated by a guillotine door. In a typical trial (acquisition trial), the animal is placed in the bright compartment and readily enters the dark compartment. At that moment, the door separating the two compartments automatically closes, and the animal receives a brief mild electric shock (0.3mA-3s). During a subsequent trial (retention trial), the latency to enter the dark compartment is recorded as an index of memory consolidation. The longer the latency to enter the dark compartment, the better the animal is supposed to remember it received an electric shock during a previous trial [3].

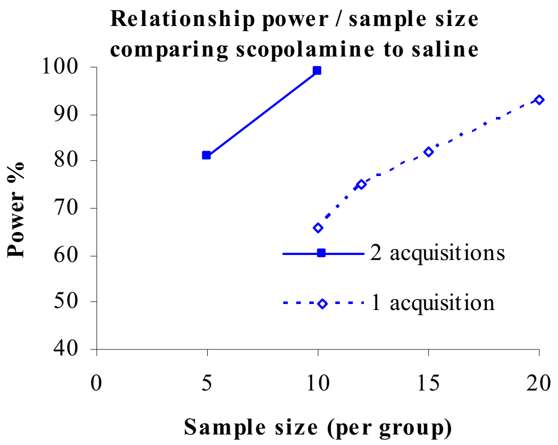
The experimental schemes found in the literature for the inhibitory avoidance paradigm substantially vary in terms of shock intensity, duration and number of trials [4]. Increasing the number of trials or increasing the sample size obviously results in a decreased throughput. The shortest protocol for an inhibitory avoidance test therefore comprises two trials: a single acquisition trial followed by a single retention trial run 24h later. The downside of this protocol is the large inter-individual variability observed in the responses, which implies that a large sample size is usually required to obtain statistically meaningful information [5,6].

In our laboratory, we optimized the 2-trial inhibitory avoidance protocol (one acquisition trial and one retention trial at 24h later) in order to test the effects of drugs against scopolamine-induced memory deficit. Scopolamine is a muscarinic receptor antagonist, which induces a cholinergic deficiency modeling to some extent the memory deficits observed in Alzheimer's disease [7]. All experiments used 8- to 9-week old male C57Black6J mice (25-30g) purchased from Charles River– France, and housed in groups of 5-6 in polypropylene cages under standard conditions (20°C, light/dark cycle 12h/12h, water and food ad libitum). They were habituated to these housing conditions for 1.5 weeks prior to experimentation. All experiments were carried out according to the European guideline 86/609/CEE and 2003/65/CE and to the Belgian legislation from August 14th, 1986 and its amendments.

We aimed at improving throughput by trying to find out a balance between a) minimizing the number of trials, and (b) maximizing the scopolamine-induced memory deficit in order to increase the window available for modulating the deficit. The 2-trial protocol was improved by adding an acquisition trial run shortly (2 minutes) after the first acquisition trial. Upon stepping through into the dark side mice received a non-escapable mild electric shock (3mA-3s). The cut-off times were 120 sec for the first acquisition trial, and 300s for the second acquisition trial as well as for the retention trial. The retention was run 24 h later as in the 2-trial protocol (see experimental scheme). This procedure was preceded 24h earlier by a session where mice were handled and habituated to the experimental room. Vehicle, test compounds and scopolamine (0.3 mg/kg) were administered i.p. 30 min. prior to the first acquisition trial. Only vehicle was administered 30 min prior to the retention trial.

The results indicated that the addition of a second acquisition trial increased the differences in latencies between vehicle and scopolamine-treated groups. In consequence, the statistical power [8], which is the probability of statistically detecting true differences between groups, was largely increased (see chart) so that the sample size needed to detect a statistically significant difference between saline and scopolamine-treated





groups was divided by three. For example, to reach a power of 80%, the sample size with a single acquisition was 15, whereas the same power was reached with only 5 mice in the two-trial acquisition protocol.

Using a 3-trial protocol, we also tested several reference drugs for their potential to reverse the scopolamine-induced memory deficit. These drugs were tacrine and donepezil, two acetylcholine esterase inhibitors, and thioperamide, an H3 receptor antagonist. The comparison of the efficacy over time showed (a) that the retention latencies for vehicle and scopolamine-treated groups were pretty stable across the year, and (b) that tacrine reversal of the scopolamine-induced memory deficit was reproducible throughout the year.

Altogether, our results indicate that the addition of a second acquisition trial shortly after the first acquisition trial increased the amplitude of the scopolamine-induced deficit, which improved the statistical power, and consequently decreased the sample size needed to show clear cut results. This protocol appears therefore to be an interesting 'compromise' between time investment (number of trials) and gain in statistical power without being forced to increase the dose of scopolamine to induce larger deficits. It also fits with the ethic's rule for animal experimentation proposing to design protocols for decreasing the number of animals used in laboratory tests.

References

1. Banfi, S., Cornelli, U., Fonio, W., Dorigotti, L. (1982), A screening method for substances potentially active on learning and memory. *J Pharmacol Methods.* **8(4)**: 255-263.
2. Kallman, M.J, Condie, L.W. Jr. (1985), A test battery for screening behavioral teratogens in mice. *Neurobehav Toxicol Teratol*. **7(6)**:727-731.
3. Graham, J.H., Buccafusco, J.J. (2001), in Buccafusc, J.J. CRC press, Methods of behavior analysis in neuroscience 141-151.
4. Sahgal, A., Passive avoidance procedures (1993), in Shagal, A., Oxford university press, Behavioral neuroscience: A practical approach 1: 49-56.
5. Elrod, K., Buccafusco, J.J. (1988), An evaluation of the mechanism of scopolamine-induced impairment in two passive avoidance protocols. *Pharmacol Biochem Behav.* **29(1)**: 15-21.
6. Lamberty, Y., Gower, A.J. (1990), Age-related changes in spontaneous behavior and learning in NMRI mice from maturity to middle age. *Physiol Behav.* **47(6)**: 1137-1144.
7. Ebert, U., Kirch, W. (1998), Scopolamine model of dementia: electroencephalogram findings and cognitive performance. *Eur J Clin Invest* **28(11)**: 944-949.
8. Hamilton, M.A., Collings, B.G. (1991), Determining the Appropriate Sample Size for Nonparametric Tests for Location Shift. *Technometrics* **33(3)**:327-337

Factors influencing isolation distress vocalization in mouse pups

P.M.Verdouw[1,2], K.Adamzek[1,2], M.J.V.van Bogaert[1,2], B.Olivier[1,2], and L.Groenink[1,2]

[1]*Utrecht Institute of Pharmaceutical Sciences, Section Psychopharmacology, Utrecht University, Utrecht, The Netherlands, P.M.Verdouw@uu.nl*

[2]*Rudolf Magnus Institute of Neuroscience, Utrecht, The Netherlands*

Rodent pups emit ultrasonic distress vocalizations (USVs) when separated from their mother and littermates. [1,2] This emotional response can be used to detect anxiolytic properties of drugs. Anxiolytic compounds reduce the duration and number of calls [3]. The ultrasonic distress vocalization test has good predictive validity in rats. With the increasing use of transgenic and knock-out mice in brain research, the ultrasonic distress vocalization test is now also more frequently performed in mice to detect differences in anxiety behavior between genotypes [4]. Here we report on factors that influence ultrasonic distress vocalization in mice. We studied differences between background strains (129Sv, Swiss Webster, C57Bl6/J) and gender, as well as the influence of weight, age and temperature of the test environment. Additionally, we measured USV in mice overexpressing corticotropin releasing factor (CRF transgenics) and studied the effect of the anxiolytic diazepam in wildtype mice and CRF transgenics. CRF is an important neuropeptide, which is released in various brain regions at times of stress. CRF integrates the behavioral and autonomic response to stress. There are indications that too much CRF may contribute to the development of anxiety disorders[5].

To conduct the test, mouse pups were put on a metal plate, which was kept at 19 or 30 °C. On top of the test plate, a chamber was mounted with a batdetector, set on 80 kHz (S-25, Ultra Sound Advice, UK) . Audio data ran through a filter to a computer running Ultravox 2.0 (Noldus Inc., Wageningen, The Netherlands). Animals remained on the plate for 5 minutes while vocalisations were registered.

An important finding of our studies was that weight is a better selection criterion to achieve optimal vocalization than age. Furthermore, we found that 129Sv emitted most USVs, followed by Swiss Webster and then C57Bl6J. This finding can be used when choosing an optimal background strain for genetic modification. CRF transgenics (C57bl6J background) vocalized significantly more than wild type mice, indicating that CRF excess in the brain enhances the anxiety response. Diazepam reduced USVs in both genotypes. Interestingly, CRF transgenics were far more sensitive to the anxiolytic effect of diazepam than there littermates, suggesting that elevated CRF levels may result in hypersensitivity of the GABA benzodiazepine receptor complex.

In conclusion, when optimizing the conditions to measure USVs in mice, this test can be used to detect genotype differences as well as differences in pharmacological sensitivity between genotypes.

References

1. Noirot E. (1966).Ultrasounds in young rodents: I. Changes with age in albino rats. *Anim Behav* **14**, 459-462.
2. Noirot E. (1972). Ultrasounds and maternal behavior in small rodents. *Dev Psychobiol* **5**, 371-387.
3. Olivier B, Molenwijk HE, van der Heyden JAM, van Oorschot R, Ronken E, Mos J, Miczek KA. (1998)a. Ultrasonic vocalisation in rat pups: effects of serotonergic ligINDS. *Neurosci Biobehav Rev* **23**, 215-227.
4. Fish EW, Sekinda M, Ferrari PF, Dirks A, Miczeck KA. (2000). Distress vocalizations in maternally separated mouse pups: modulation via 5-$HT_{1A,}$ 5-HT_{1B} and $GABA_A$ receptors. *Psychopharm* **149**, 277-285.
5. Arborelius L, Owens M J, Plotsky P M and Nemeroff C B. (1999). The role of corticotropin-releasing factor in depression and anxiety disorders. *Journal of Endocrinology* **160**, 1-12.

Comparison of the effects of antipsychotic drugs in two antipsychotic screening assays: swim-induced grooming and apomorphin climbing test in mice

R. Kedves, K. Sághy, and I. Gyertyán
Behavioural Pharmacology, Gedeon Richter Ltd, Budapest, Hungary, r.kedves@richter.hu

Introduction

The antagonism of apomorphine-induced climbing behaviour and sniffing and the swim-induced grooming test are widely used models for screening antipsychotic drugs. It is hypothesized that blockade of climbing behaviour evoked by apomorphine is related to the dopamine receptor blockade in the nucleus accumbens and antagonism of stereotyped sniffing produced by apomorphine is related to the blockade of dopamine D2 receptors in the striatum [1]. The dopamine receptor antagonist antipsychotics inhibit both apomorphine-induced behaviours, although compounds with strong 5-HT2A inhibitory effect show higher potency in blocking the climbing vs. the sniffing response. The grooming behaviour induced by immersion in water involves mainly dopamine D1 receptors [2]. This behaviour is inhibited in a dose dependent manner by dopamine receptor antagonists. [3]. The aim of our study was to compare these two screen assays in terms of their sensitivity to the actions of various antipsychotic compounds. We investigated the effect of a conventional neuroleptic (haloperidol), atypical antipsychotics (clozapine, olanzapine, risperidone, ziprasidone, amisulpride) and new generation, partial D2 agonist antipsychotics (arpiprazole, bifeprunox) in these screening methods.

Methods

Climbing and sniffing Male CD-1 mice were injected orally with the test compound or vehicle. Fifty minutes later the animals were placed into cylindrical cages, with walls of vertical metal bars 2 mm diameter 1 cm apart, surmounted by a smooth surface. After 10 minutes habituation the mice were treated with apomorphine (1.5 mg/kg sc.). Ten minutes after APO treatment each animal was observed for 15 minutes. Every minute the climbing behaviour was scored as follow: four paws on the floor (0), forefeet grasping the wall (1), four paws grasping the wall (2). Animals were also rated for repetitive sniffing as a measure of stereotypy according to the following scale: 0 = no sniffing, 1 = moderate sniffing, little snout contact with cage walls or floor, 2 = constant sniffing, persistent snout contact. Scores for both behaviours were summed for each individual and group means were calculated. Drug effect was expressed and plotted as percentage inhibition of the apomorphine-induced behaviour. Dose-response curve was plotted for each compound and ED_{50} values were determined by linear regression.

Swim incuced grooming in mice Male NMRI mice were used. Sixty minutes after drug treatment, mice were placed individually in swimming cylinders filled with waterfor 3 min. They were then removed and dried with towel for 30 sec. and placed immediately into single perspex boxes. The number and the total duration of grooming episodes in seconds were recorded for 15 min. Dose-response curve was plotted for each compound and ED_{50} values were determined by linear regression.

Results

ED_{50} values of antipsychotics in apomorphine-induced climbing and sniffing and swim-induced grooming test are summarized in Table 1. Apomorphine-induced climbing and sniffing were potently and dose-dependently blocked by the compounds. The order of potency of antipsychotics on climbing inhibition was bifeprunox = risperidone ~ haloperidol > aripiprazole ~ olanzapine > ziprasidone ~ clozapine >> amisulpride. The order of potency of antipsychotics in blocking sniffing was bifeprunox > haloperidol ~ risperidone > aripiprazole > olanzapine = ziprasidone >> amisulpride > clozapine.

Post-swimming grooming time was inhibited dose-dependently by risperidone = bifeprunox > haloperidol > olanzapine > clozapine > aripiprazole = ziprasidone. Interestingly, amisulpride elevated rather than decreased the grooming time at doses of 30 and 60 mg/kg ip while it had no effect after 120 mg/kg ip.

Table 1. *Oral ED_{50} values of antipsychotics in the apomorphine antagonism and swim-induced grooming test. Values are in mg/kg.*

	Apomorphine-induced		Swim-induced
	climbing	sniffing	grooming time
Haloperidol	0.17	0.32	0.65
Clozapine	5.3	18	2.1
Olanzapine	1.3	4.4	1.2
Risperidone	0.12	0.45	0.30
Aripiprazole	0.97	2.2	5.1
Bifeprunox	0.1	0.14	0.38
Ziprasidone	3.9	4.6	5.4
Amisulpride (ip.)	30	12	increase

Conclusion

Compound which showed potent apomorphine-induced climbing inhibition relative to sniffing inhibition (risperidone, olamzapine, clozapine) have high 5-HT_{2A} receptor antagonist effect. Olanzapine and clozapine, which have considerable dopamine D1 receptor antagonist effect, showed potent blockade of swim-induced grooming relative to their climbing inhibition.

As the two widely used screening methods showed different sensitivity to the various antipsychotics their parallel use in antipsychotic screening systems is therefore warranted.

References

1. Moore, N.C. and Gershon S. (1989). Which atypical antipsychotics are identified by screening tests? *Clinical neuropharmacology* **12(3):** 167-184
2. Van Wimersma Greidanus, T. B., Maigret, C., Torn, M., Ronner, E., Van Der Kracht, S., Van Der Wee, N. J. A. and Versteeg, D. H. G. (1989). Dopamine D-1 and D-2 receptor agonists and antagonists and neuropeptide-induced excessive grooming. *European journal of pharmacology* **173**: 227–231
3. Chesher GB, Jackson DM. (1980). Post-swim grooming in mice inhibited by dopamine receptor antagonists and by cannabinoids. *Pharmacology, biochemistry, and behavior* **13(3):** 479-81
4. Chesher GB, Jackson DM. (1981). Swim-induced grooming in mice is mediated by a dopaminergic substrate. *Journal of neural transmission* **50(1)**: 47-55

Automated distinguishing of mouse behavior in new environment and under amphetamine using decision trees

A. Konushin[1], E. Lomakina-Rumyantseva[1], D. Kropotov[2], D. Vetrov[1], A. Cherepov[3], and K. Anokhin[3]
[1]*Graphics and multimedia lab, Moscow State University, Moscow, Russia. {ktosh, vetrovd}@graphics.cs.msu.ru, lr2kate@gmail.com*
[2]*Dorodnicyn Computing Center of the Russian Academy of Sciences, Moscow, Russia, dkropotov@yandex.ru*
[3]Department of Systemogenesis, *P.K. Anokhin Institute of Normal Physiology, Moscow, Russia, k_anokhin@yahoo.com*

Abstract

Traditional activity measures do not provide a clear discrimination between mouse behavior in novel environment or under various psychomotor stimulants like *d*-amphetamine [1]. We propose a new approach based on machine learning. A decision tree classifier is trained on a set of mouse trajectories. Only x- and y- coordinates are currently used. After classifier is trained, the classification can be applied in real-time with respect to the output of video tracking system during experiment. The developed approach has shown promising results on distinguishing effects of new environment and d-amphetamine according to mouse behaviors.

Keywords

Video tracking, decision trees, novel environment, amphetamine, machine-learning.

Introduction

In recent years automated home cage systems have been introduced as a solution to the need of high-throughput behavior screening procedures [2]. Many of such systems are equipped with video tracking modules. However, traditional activity measures cannot automatically distinguish the effects of different drugs and environment, leaving the task to the human observer. This significantly limits the throughput capacity of such systems, because automated cage systems provide information 24 hours a day, generating an enormous volume of data.

Several new activity-based measures were recently proposed, e.g. [3], but they are based on pure observation and heuristically constructed. We propose a novel approach based on machine learning. This approach gives us a general framework for automatic development of classifiers for distinguishing the effects of various drugs and stimuli.

Proposed method and experiments

We had three types of recorded mouse behavior: mouse in a new environment (class 1), under effect of d-amphetamine (5 mg/kg) (class 2) and in familiar environment (class 3). Totally, we had 15 mice and 15 minutes (25000 time samples) behavior record of each, 6 mice from class 1 and class 2, and 3 mice from class 3. In order to increase data set we split each path into four equal parts thus obtaining enough representative samples (60 in total).

To design classification algorithm we computed a number of features according to mouse locomotor motion in the cage. After several tests the following three key features of path were identified: mean speed (v_mean), speed standard deviation (v_std), and correlation coefficient between speed and acceleration. The latter was computed in the following way:

$$\text{Corr\_coeff} = \frac{1}{n-4}\sum_{i=3}^{n-1}(v_i - Ev)(a_i - Ea)\Big/\sqrt{DvDa}\,,$$

where $v_i = \sqrt{(x_i - x_{i-1})^2 + (y_i - y_{i-1})^2}$ - speed at the i-th time point, $a_i = v_i - v_{i-1}$ – acceleration at the *i*-th time point, $Ev = \frac{1}{n-1}\sum_{i=2}^{n} v_i,\ Ea = \frac{1}{n-2}\sum_{i=3}^{n} a_i,\ Dv = \frac{1}{n-2}\sum_{i=2}^{n}(v_i - Ev)^2,\ Da = \frac{1}{n-3}\sum_{i=3}^{n}(a_i - Ea)^2$.

We used decision trees CART algorithm [4] which gave us the following decision rule:

If ((corr_coeff <= 0.51425) **and** (v_mean <= 0.00245))
Then prediction = class 3 (probability = 1.000000)
Else if ((corr_coeff <= 0.51425) **and** (v_mean > 0.00245) **and** (v_std <= 0.00805))
Then prediction = class 1 (probability = 1.000000)
Else if ((corr_coeff <= 0.51425) **and** (v_mean > 0.00245) **and** (v_std > 0.00805))
Then prediction = class 2 (probability = 1.000000)
Else if (corr_coeff > 0.51425)
Then prediction = class 2 (probability = 0.869565)

The right classification in leave-one-out mode was achieved in 95% cases (57 of 60). This result proves that the behavioral features mentioned above discriminate reasonably well these states of mouse.

Note that the features used can be computed in real-time by using only part of mouse trajectory up to the current moment. This makes it possible designing on-line version of classifier capable to estimate the probability of each class at every time moment. Let x_j be the number of time samples at which the mouse was classified to class *j*. Then the probability that mouse belongs to given class *k* can be computed as $p_k = (x_k + c)/(x_1 + x_2 + x_3 + 3c)$. Here the constant c serves for preventing large fluctuations of probability in the beginning of trajectory. In our experiments we put $c = 50$. Figure 1 shows the typical changes of probabilities which are computed for each time sample. Using diagrams similar to figure 1, we can observe classifier's performance and make conclusion in a real-time mode.

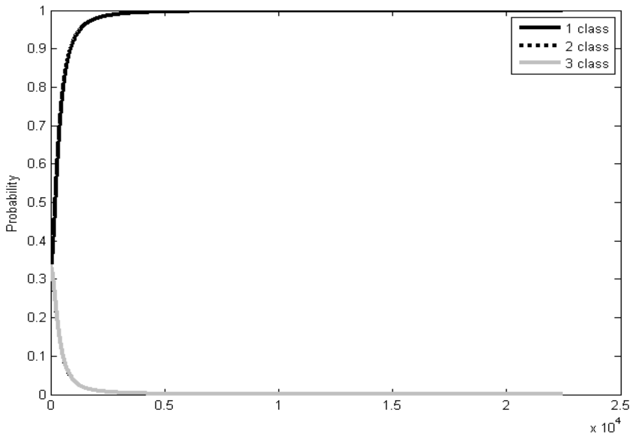


Figure 1. *Probabilities of different classes for particular mouse for all time moments.*

Conclusion and future work

Our proposed method has shown promising results for on-line classification of distinguishing the effects of amphetamine and new environment. We plan to introduce new features based on nose and tail base point tracking, and automatically labeled behavioral acts. Also on-line classification will be tested on long experiments for recognition of different doses of d-amphetamine and vanishing of drug or other stimuli effect.

References

1. Cherepov A.B., Lukashev A.O., Novosyolov I.A., Rayevsky K.S., Anokhin K.V. (2002) Relationships between novelty- versus d-amphetamine- induced behavior and brain c-fos expression in two inbred mouse strains: a functional systems analysis, *Measuring Behavior* 2002.
2. Spruijt B.M., DeVisser L.(2006) Advanced behavioral screening: automated home cage ethology. *Drug Discovery Today: Technologies*., **3(2)**, 231-237.
3. Kafkafi.N, Elmer G.I. (2005) Activity density in the open field: a measure for differentiating the effect of psychostimulants. *Pharmacology, Biochemistry and Behavior*, **80**, 239–249.
4. Breiman L., Friedman J., Olshen R., Stone C. (1984) Classification and Regression Trees. *Belmont, California*.

Active allothetic place avoidance task in study of cognitive disturbance in animal model of schizophrenia in rats

Lukas Rambousek[1,2], Karel Vales[1], Jana Kenney[1], and Ales Stuchlik[1]

[1]*Department of Neurophysiology of Memory and Computational Neuroscience, Institute of Physiology, Academy of Sciences, Prague, Czech Republic, rambosuek@biomed.cas.cz*

[2]*Department of Organic Technology, Institute of Chemical Technology Prague, Czech Republic*

Since the ethiopathology of schizophrenia is not fully understood, it is difficult to establish its animal model with full construct validity. Recently applied models are based on the central blockade of glutamate receptors, namely their NMDA subtype. It has been observed that administration of compounds that block NMDA receptor-dependent neurotransmission (phencyclidine, ketamine, dizocilpine) elicits a psychotic state when applied to healthy humans and worsens the psychotic symptoms when administered to schizophrenic patients [1]. Application of high-affinity non-competitive NMDA-receptor antagonist MK-801 (dizocilpine) was proposed as an animal model of schizophrenia which proved to have relatively high predictive validity [2]. Animals treated with MK-801 exhibit typical changes in behavior including hyperactivity, stereotypic behaviors, defective habituation, impaired attention, simpler behavioral repertoire and general behavioral primitivization [3].

Testing the cognitive abilities of animals with experimentally induced psychotomimetic state requires specific behavioral paradigms, which should be dry land-based (more natural than the Morris water maze) and, what is more important, they should have a relatively high cognitive demand for their efficient solution [4]. Morris water maze or radial maze focus on testing allocentric memory processes and they do not sufficiently cover other cognitive functions. By contrast, the Active Allothetic Place Avoidance test (AAPA) is useful for detecting cognitive disorganization. We propose that AAPA requires more complex cognitive abilities than classical behavioral paradigms. The principle of function of the place avoidance tasks is that rats are moving on a uniform circular arena, on which an arbitrarily located unmarked sector is defined, entering of which is punished by a mild footshock.

Briefly, AAPA setup consists of a smooth metallic circular arena (80 cm in diameter) enclosed with a 30 cm high transparent Plexiglas wall and elevated 1m above the floor of the experimental room containing many visual landmarks. At the beginning of each training session, a rat was placed on the rotating arena (1 rpm), where a directly imperceptible 60◦ to-be-avoided sector (shock sector) was defined by the custom-based computerized tracking system, located in an adjacent room. The location of the shock sector could be determined exclusively by its spatial relations to distal orienting cues located in the room. The rat wore an infrared light-emitting diode (LED) fixed between its shoulders with a light latex harness, and its position was tracked every 40 ms and recorded onto a computer track file, allowing subsequent reconstruction of the track with an off-line analysis program (TrackAnalysis, Biosignal Group, USA). Whenever a rat entered the shock sector for more than 0.5 s, mild electric shocks (50 Hz, 0.5 s) were delivered in intervals of 1.5 s until the rat left the shock sector for at least 0.5 s. The shocks were delivered through a thin subcutaneous low-impedance wire implant on the back of the rat standing on the grounded floor. The appropriate shock current (ranging between 0.2 and 0.7 mA) was individualized for each rat to elicit a rapid escape reaction but to prevent freezing. Since, the arena was rotating; the rat had to move actively away from the shock in the direction opposite to the arena rotation, otherwise it was passively transported to the shock sector. Experimental sessions in AAPA lasted 20 min and each rat had one session every day, carried out during daylight hours.

Successful performance of the task requires that the animal identifies its position in the room frame and also the position of the shock sector in the room frame. Since the arena-frame information (droppings, urine and scent marks self-generated by rats) rotates with respect to the shock sector the animal must also recognize these arena-based cues as distinct from the stimuli that come from the room frame and ignore the irrelevant arena-frame for localization of the shock sector. Such an ability was described to be hippocampus-dependent and called "cognitive coordination". This requirement that the subject differentiates between relevant and irrelevant stimuli is similar to the concept that schizophrenic patients are often unable to differentiate between relevant and irrelevant stimuli because their information processing is impaired.

This research was supported by grants IGA-MH-NR/9178-3, GACR-309/07/0341, and MEYS-CR-project-1M0517.

References

1. Coyle JT, Tsai G, Goff D. (2004). Converging evidence of NMDA receptor hypofunction in the pathophysiology of schizophrenia. *Ann. NY Acad. Sci.,* **1003**, 318-327.
2. Farber N.B. (2003). The NMDA-receptor hypofunction model of psychosis, *Ann. N.Y. Acad. Sci.*, **1003**, 119–130.
3. Koek W, Woods JH, Winger GD. (1988) MK-801, a proposed noncompetitive antagonist of excitatory amino acid neurotransmission, produces phencyclidine-like behavioral effects in pigeons, rats and rhesus monkeys. *J. Pharmacol. Exp. Ther.*, **245**, 969-974.
4. Stuchlik A., Rezacova L, Vales K, Bubenikova V, Kubik S. (2004). Application of a novel Active Allothetic Place Avoidance task (AAPA) in testing a pharmacological model of psychosis in rats: comparison with the Morris Water Maze. *Neurosci. Letters*, **366**, 162-166.

Autistic-like behaviour in the Parvalbumin knockout mouse

P.J.M. Gregory[1], D. Wolfer[2,3], I. Drescher-Lindh[2], H. Welzl[2,3], and B. Schwaller[1]
[1]*Institute of Anatomy, University of Fribourg, Fribourg, Switzerland, pgregory@unifr.ch*
[2]*Institute of Anatomy, University of Zürich, Zürich, Switzerland, david@dpwolfer.ch*
[3]*Department of Biology, ETH, Zürich, Zürich, Switzerland*

Introduction

The EF-hand calcium-binding protein, parvalbumin (PV), is a cytosolic protein expressed in a subset of GABA-ergic interneurons of the cerebral cortex, Purkinje, stellate and basket cells of the cerebellum, fast-twitch skeletal muscle fibres and epithelial cells of the kidney distal nephron. It functions primarily as a mobile, slow acting calcium buffer and plays a principal role in intracellular calcium homeostasis. Studies on the PV knockout mouse (PV-/-) showed an outwardly normal phenotype and lifespan. However, skeletal muscle fibres showed a reduced relaxation rate and increased fatigue resistance, coupled to a ~50 % increase in mitochondrial volume [1]. This increase in mitochondria was also observed in the soma of Purkinje cells in PV-/- mice [2]. The autism spectrum disorders (ASD) are human neuro-developmental disorders, characterized primarily by impairments in social interaction / social communication and repetitive, restrictive behaviour. Autism is increasingly being seen as an impairment or imbalance of the GABA-ergic system and due to the selective expression pattern of PV, it was hypothesized that PV-deficiency would correlate with autistic-like behaviour.

Results

Male PV-/- and wildtype (WT) mice were investigated for motor coordination, locomotion, anxiety, nociception, pre-pulse inhibition (PPI), social interaction / social novelty and restrictive, repetitive behaviour. PV-/- mice were indistinguishable from WT in tests for anxiety-like behaviour and exhibited no adverse behaviour in any of the anxiety-related assays. However, even though their average speed and distance travelled in an open-field assay were no different from WT controls, they showed significantly greater micro-linearity and fewer rearing events, suggesting a stereotypic form of locomotion. They also performed less horizontal turns (and therefore reduced “escape behaviour”) in a rotarod assay, despite performing equally well, if not better, than WT mice on 3 different protocols. Both male and female PV-/- showed reduced pain sensitivity in hotplate and tail-flick assays compared to WT mice. PV-/- also showed reduced startle response at higher decibel levels than WT mice and a slightly, though not significantly, reduced (PPI). In a T-maze reversal assay, PV-/- mice showed a significantly reduced ability to reverse from an acquired behaviour compared to controls, but in an adapted 3-chamber assay for social interaction and social novelty, compared to WT the PV-/- mice showed no preference for an inanimate object (empty cage) to a novel stranger mouse, or between the 1st stranger and a 2nd mouse introduced later.

Conclusion

Deficiencies in social interaction, social communication and repetitive, restrictive behaviour are the key diagnostic criteria for ASD in humans. Adapting existing mouse behavioural assays to test for these behaviours is key to modelling autistic-like behaviour in rodents. In a series of experiments to screen for such behaviour, PV-/- exhibited deficits in reversal learning (i.e. repetitive, restrictive behaviour), but not in social interaction or social novelty. In further tests to screen for symptoms associated with, though not clinically diagnostic of, ASD, the PV-/- mice exhibited motor stereotypy and significantly reduced pain sensitivity. Even though PV deficiency may not be the underlying genetic aetiology of ASD, its absence may well be manifest in some of its symptoms. This correlates well with the documented reduction in PV-positive neurons in patients with other neurological disorders involving GABA-ergic transmission (schizophrenia, bipolar disorder) [3] and other transgenic mouse models of autism [4]. The association between chromosome 15q disorders (Prader-Willi Syndrome, Angelman syndrome), GABA receptor expression and autism is also well documented.

References

1. Schwaller B, Dick J, Dhoot G, Carroll S, Vrbova G, Nicotera P, Pette D, Wyss A, Bluethmann H, Hunziker W, Celio MR. (1999) Prolonged contraction-relaxation cycle of fast-twitch muscles in parvalbumin knockout mice. *American Journal of Physiology,* **276**, C395-403
2. Chen G, Racay P, Bichet S, Celio MR, Eggli P, Schwaller B. (2006) Deficiency in parvalbumin, but not in calbindin D-28k upregulates mitochondrial volume and decreases smooth endoplasmic reticulum surface selectively in a peripheral, subplasmalemmal region in the soma of Purkinje cells. *Neuroscience,* **142(1)**, 97-105
3. Beasley C.L. and Reynolds G.P. (1997). Parvalbumin-immunoreactive neurons are reduced in the prefrontal cortex of schizophrenics. *Schizophrenia Research,* **24 (3),** 349-355
4. Sadakata T, Washida M, Iwayama Y, Shoji S, Sato Y, Ohkura T, Katoh-Semba R, Nakajima M, Sekine Y, Tanaka M, Nakamura K, Iwata Y, Tsuchiya KJ, Mori N, Detera-Wadleigh SD, Ichikawa H, Itohara S, Yoshikawa T, Furuichi T. (2007). Autistic-like phenotypes in Cadps2-knockout mice and aberrant CADPS2 splicing in autistic patients. *Journal of Clinical Investigation,* **117**, 931-943

Behavioral phenotyping of a murine AD-model in a semi-naturalistic environment using RFID tracking

L. Lewejohann, N. Reefmann, A.M. Hoppmann, and N. Sachser
Department of Behavioral Biology, University of Muenster, Germany, www.ethologie.de, Ljohann@phenotyping.com

Neurodegenerative disorders such as Alzheimer's disease (AD), which are characterized by an advancing cognitive decline are increasingly threatening public health. Most animal models of AD consist of transgenic mice that are usually housed singly or in unisexual groups in small barren cages. Such restricted environments, however, prevent the mice from showing a variety of species-specific behavior and consequently may constrain behavioral phenotyping. The aim of this project was to characterize TgCRND8-mice carrying a genetic disposition (human $APP_{Swe+Ind}$) to develop Alzheimer-like pathology and their wild-type conspecifics in a semi-naturalistic environment (SNE). The SNE measured 1.75 by 1.75 by 2.1m (L x W x H) and contained several floors (see Figure 1). The population comprised male and female mice of both genotypes and was allowed to grow to a size of 40 adults. Mice were individually marked using subcutaneously injected RFID transponders as well as a color coding scheme on their tails and ears. In a first step behavioral observations at different ages of the mice were conducted by thoroughly trained experimenters distinguishing up to 55 unique behavioral patterns from various behavioral domains. First results revealed surprisingly little significant differences between genotypes that were true at all ages and for both sexes.

Interestingly the mice established a complex social structure comprising several territories held by dominant males. In a second step direct observations were complemented by RFID-antennas placed at strategically chosen spots within the SNE. The transponder ID is read while a mouse traverses the electromagnetic field which is established by the ring antennas, e.g. when passing through tubes or visiting drinking places. A software that was developed for this setup allowed constant monitoring of several locomotory, activity, and behavioral patterns on a 24h/7d basis. The automated RFID-tracking system was evaluated by direct observation of movement data that correlated highly significantly with the number of antenna contacts. Dominant males patrol their territory borders more frequently than subdominants and thereby triggered significantly more antennae contacts than subdominant animals. This indicates that the RFID-system applied here may indeed facilitate behavioral observations as it allowed the determination of social status by the number of antenna contacts. Both, transgenic and wild-type mice were able to achieve high dominance positions and were able to hold a territory. To test spatial memory performance, a modified Barnes Maze was conducted within the SNE. This test revealed that deficits that are known from TgCRND8 mice are still present in the SNE. However, marked differences in activity and stereotypic behavior that were observed in TgCRND8 mice in standard cages [1] could not be observed in the SNE. This indicates, that a physically and socially enriched environment can modify the way an individual is able to cope with the disease.

Figure 1. *Semi-naturalistic environment (SNE). RFID marked mice were are automatically tracked when passing a ring-antenna.*

References

1. Ambrée, O., Touma, C., Görtz, N., Keyvani, K., Paulus, W., Palme, R., Sachser, N. (2006). Activity changes and marked stereotypic behavior precede Abeta pathology in TgCRND8 Alzheimer mice. *Neurobiology of Aging*, **27**, 955-964.

Transgenic overexpression of MMP-9 in rats results learning and memory alterations

M. Wawrzyniak[1], A. Kiryk[1], V. Lioudyno[1], K.Z. Meyza[2], D. Owczarek[1] and L. Kaczmarek[1]
[1]*Department of Molecular and Cellular Neurobiology, Nencki Institute of Experimental Biology, PAS, Warsaw, Poland, m.wawrzyniak@nencki.gov.pl*
[2]*Department of Neurophysiology, Nencki Institute of Experimental Biology, PAS, Warsaw, Poland*

Matrix metalloproteinases are a major group of enzymes regulating cell-matrix composition, which are essential for many biological phenomena. Matrix Metalloproteinase-9 (MMP-9) has recently emerged as an important molecule in control of extracellular proteolysis in synaptic plasticity. Using conventional transgenesis, we have created rats with overexpression of MMP-9 limited to neurons (MMP-9-gene is under control of synapsin-1 promoter).

The animals, along their wild-type siblings, were exposed to a variety of behavioral tasks. To analyze anxiety, we employed Elevated Plus Maze (EPM) test [1, with modifications] that showed 12-15-months old transgenic rats spending more time in the center zone, while the wild type rats stayed in the closed arms (none of the rats entered the open arms, see Figure 1A, B). Lack of entrance to the open arms was unexpected, and prompted us to repeat the test on 3-months old animals. The young rats were entering the open arms, and no significant difference between transgenic and control group was observed (see Figure 1C, D). Boguszewski and Zagrodzka [2] as well as Bessa et al. [3] reported similar age-related decline in entering the open arms. However, in their experiments the old animals still entered the open arms. This apparent discrepancy can be explained by differences in the experimental set-up employed. We used white curtains hanging around and near plus maze (rats oriented toward environment without extra-apparatus cues) and center light which might be a mild stressor for the animals. The other authors used dim red light or fluorescent lamps mounted above the maze so that all arms were equally illuminated, so their light conditions were quite different.

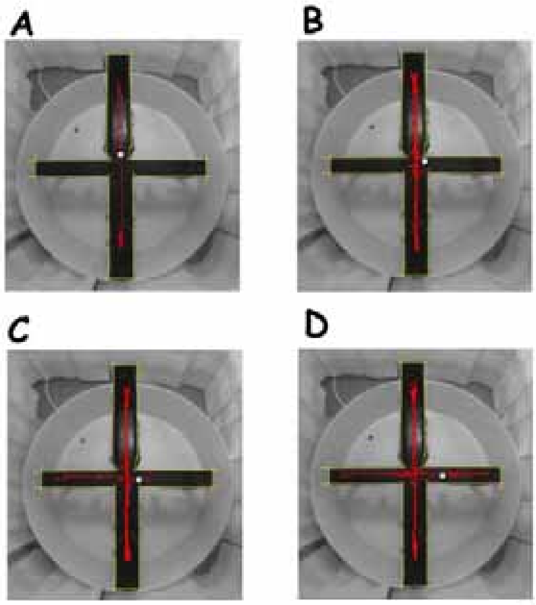


Figure 1. *Exemplary records of rat's footpath in Plus Maze recorded by Ethovision. A – old wild type rat, B – old transgenic rat, C – young wild type rat, D – young transgenic rat. In Plus Maze test old transgenic rats were more active than control rats and none of old rats entered to the open arms.*

In additional experiments, we noticed differences in general locomotor activity between old transgenic vs. wild type rats. In the Open Field (OF) analyzed by Ethovision, no significant differences between young transgenic and control rats were revealed. However, the old transgenic rats were more active in OF (see Figure 2A, B) and, in addition, performed better on a RotaRod than controls.

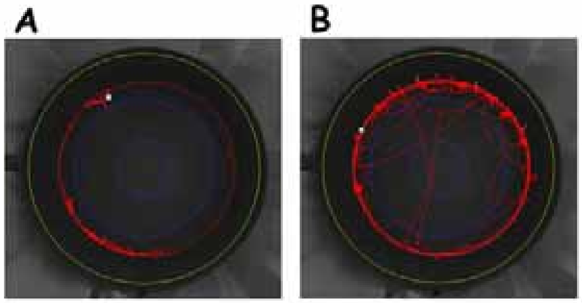


Figure 2. *Exemplary records of rat's footpath in Open Field recorded by Ethovision. A – old wild type rat, B – old transgenic rat. Transgenic rats were more active than control in Open Field test.*

We have also carried out a Flavor Preference (FP) test that is connected with appetitive learning involving flavor-reward associations [4, with modifications]. During the first 4 days of the experiment each rat was presented in the morning with water for 15 min, followed 5 hrs later by 1 hr water exposure. Then, the animals were divided into two groups. The first one was exposed to sweet orange and non-sweet apple juice on every second day in the morning and the second group to non-sweet orange and sweet apple juice under the same time regime. The juices were presented in dimmed bottles to avoid recognizing the color of the liquid. All the animals were exposed to water in the afternoon, as before. On the days 11-14, rats had a choice between non-sweet orange and non-sweet apple juice. The amount (ml) consumed of each solution was recorded daily and was used to measure % of consumed juice with positive flavor-association (that is memory connected with flavor). In this experiment we showed, that the old transgenic rats had a better taste preference than their wild type siblings. Rentiera et al. [5] showed that rats develop preference for the sweetened juice (drink more) and that the ability to form flavor-reward associations declines with age, resulting in impaired conditioned flavor preference, as we showed.

These results strongly suggest that MMP-9 may have an important role in control of the behavior of animals.

References

1. Pellow, S., Chopin, P., File, S.E., Briley, M. (1985). Validation of open:closed arm entries in an elevated plus-maze as a measure of anxiety in the rat. *J Neurosci Methods*., **14**, 149-167.

2. Boguszewski, P., Zagrodzka, J. (2002). Emotional changes related to age in rats--a behavioral analysis. *Behav. Brain Res.*, **18**, 323-332.
3. Bessa, J.M., Oliveira, M., Cerqueira, J.J., Almeida, O.F., Sousa, N. (2005). Age-related qualitative shift in emotional behaviour: paradoxical findings after re-exposure of rats in the elevated-plus maze. *Behav. Brain Res.*, **162**, 135-142
4. Gilbert, P.E., Campbell, A., Kesner, R.P. (2003). The role of the amygdala in conditioned flavor preference. *Neurobiol. Learn. Mem.*, **79**, 118-121.
5. Renteria, A.F., Silbaugh, B.C., Tolentino, J.C., Gilbert, P.E. (2008). Age-related changes in conditioned flavor preference in rats. *Behav. Brain Res.*, 188, 56-6.

Wild genius - domestic fool?
Spatial learning abilities of *Cavia aperea* and *Cavia aperea* f. *porcellus*

Thorsten Pickel, Sylvia Kaiser, Lars Lewejohann, and Norbert Sachser
Department of Behavioural Biology, University of Muenster, Germany, thorsten.pickel@gmx.de

Domestic animals and their wild relatives differ in a wide variety of aspects. The process of domestication of guinea pigs (*Cavia aperea* f. *porcellus*), starting 3000 to 6000 years ago, led to many changes in anatomy, physiology and behaviour of the animals compared to their wild ancestor *Cavia aperea*. Apart from the obvious changes in body size, fur colour or vocalization, one characteristic domestic trait is a reduction in brain size, which in guinea pigs accounts for about 13% in relation to body size. This might lead to lower learning abilities in domestic species compared to their wild relatives.

This study explored potential differences in learning abilities of wild and domestic guinea pigs. Therefore, animals of both forms and sexes were tested in the Morris water maze. For the wild guinea pigs, this was the first time that this spatial learning paradigm was applied. Thus, a second aim of this study was to establish this behavioural test for this species.

The maze consisted of a pool of 160 centimetres in diameter and a platform of 20 centimetres in diameter made of hyaline acrylic glass. The platform was invisible reaching two centimetres below the water surface. In the test version used, each animal was subjected to ten swimming trials within five consecutive days. In each trial of this acquisition phase the animal had a maximum time of 45 seconds to reach the hidden platform. If it did not find the platform to escape from the water, the guinea pig was put manually on it and had to stay there for 15 seconds. Between the two trials of each day the guinea pig had a recovery time of five minutes in their home cages so that their fur could dry.

Learning curves concerning distance swum as well as latency to reach the platform were analyzed. In addition, the integrals below the learning curves were compared between wild and domestic form. Subsequently, in a probe trial of 60 seconds duration without the platform present, the time the animals stayed in the correct quadrant served as an additional measure of learning success.

Significant differences in spatial learning were found between domestic and wild guinea pigs. Although all animals showed spatial learning in the Morris water maze, male and female domestic guinea pigs proved to be better learners than their wild relatives: On average, they swam a shorter distance and had also a shorter latency to reach the hidden platform. Furthermore, in the probe trial, they also spent significantly more time in the correct quadrant of the water basin.

These results demonstrate that the applied test version of the Morris water maze is a useful tool to assess spatial learning in domestic guinea pigs, and in wild guinea pigs as well. The comparatively lower learning success of the wild species might also be a result of a reduced ability to cope with the man-made environment and handling procedure. Contrary to expectations derived from the lower brain size, this study shows that the domestic form of the guinea pig does not at all perform worse than its wild ancestor. Hence, artificial selection and breeding did not lead to degenerated domestic animals with impaired cognitive abilities.

Scopolamine-induced changes in activity measured in a home cage observation system

Serena Deiana, Bettina Platt, and Gernot Riedel
School of Medical Sciences, University of Aberdeen, Scotland, UK, bms264@abdn.ac.uk

Background

The internal body clock (suprachiasmatic nucleus of the basal hypothalamus) regulates the circadian rhythms which controls sleep/awake cycles, locomotor activity and feeding/drinking behaviour in both humans and animals [1]. Various drug treatments and neurotransmitters can interfere with these rhythms [2] and consequently drug-induced effects obtained in memory experiments may be due to the direct action of the drug on circadian biological rhythms. Moreover, the effects of the drug may depend on the administration time during the 24 hr cycle.

Scopolamine, a competitive M_1 muscarinic cholinergic receptor antagonist, is known to induce cognitive deficits [3], as well as effects on locomotor activity [4, 5] and alteration of sleep stages [6].

Aim

The present study aimed to investigate the effects of sub-chronic administration of scopolamine on locomotor activity and circadian rhythms of freely moving mice using the PhenoTyper system (Noldus)

Methods

C57BL6/j male mice were habituated for two days to the PhenoTyper cage to bring the activity to basal levels. Mice were maintained on a 12:12 hr light/dark cycle (light off at 7pm). On days 3-6 mice were intraperitoneally injected for four consecutive days with either saline or scopolamine (0.5 mg/kg) 1-2 hr prior the beginning of the dark phase and returned into the PhenoTyper cage. Locomotor activity was continuously recorded (24 hr/day) with the tracking software Ethovision (Noldus), and data were analysed using time intervals of 1 hr. Further analysis of activity in light and dark phases and 3 hr post-injection was performed. Data were analysed with repeated measures Two-way ANOVA, with treatment and time point as independent variables followed by Bonferroni post-hoc test with no correction for large number of possible comparison. Thereafter data recorded during the first 3 hr post injection were averaged for each day of treatment. The null hypothesis "does scopolamine has a lower activity compared to controls" was defined and Student t-tests were used to compare the two treatment groups at each day. For all comparisons, a 95% confidence level ($p<0.05$) was set for the differences to be considered as significant.

Results

The PhenoTyper detected circadian rhythm of mice as changes in locomotor activity (see Figure 1a), with no drug-induced effects in scopolamine-treated subjects, which showed similar circadian rhythms of activity to controls ($p>0.05$). However, further analysis revealed that scopolamine induced a significant reduction in overall locomotor activity during the first three hours post-injection, this effect was only evident following the first two days of treatment with effects no further observed following subsequent administration (see Figure 1b).

Conclusions

The PhenoTyper measured circadian rhythms and locomotor activity of the mice. Scopolamine did not affect such rhythms; however, it decreased the activity during the light/dark change event (or crepuscule), during which an increase in activity is normally expected. Moreover, the decrease of locomotor activity was time-dependent with effects only observed following initial treatment, suggesting a possible tolerance to the effect of scopolamine. Scopolamine is known to increase locomotor activity [4, 5], yet the results presented in this study, suggest that such effect may depend on the time of the light-dark cycle at which the drug is administered. In conclusion, the present data indicate the versatility of the PhenoTyper as a device to investigate not only drug-induced changes in locomotor activity, and cycle-dependent effects, but also the possible establishment of behavioural tolerance to such effects of the drugs.

References

1. Weaver, D.R. (1991). The suprachiasmatic nucleus: a 25-year retrospective. *J. Biol. Rhythms*., **13(2)**, 100-12.
2. Labrecque, G., Bélanger, P.M. (1991). Biological rhythms in the absorption, distribution, metabolism and excretion of drugs. *Pharmacol. Ther*., **52(1)**, 95-107.
3. von Linstow Roloff, E., Harbaran, D., Micheau, J., Platt, B., Riedel, G., (2007). Dissociation of cholinergic function in spatial and procedural learning in rats. *Neuroscience*, 146**(3)**, 875-889.
4. Shannon, H.E., Peters, S.C. (1990). A comparison of the effects of cholinergic and dopaminergic agents on scopolamine-induced hyperactivity in mice. *J. Pharmacol. Exp. Ther*., **255(2)**, 549-53. 675.
5. Toide, K. (1989). Effects of scopolamine on extracellular acetylcholine and choline levels and on spontaneous motor activity in freely moving rats measured by brain dialysis. *Pharmacol. Biochem. Behav*., **33(1)**, 109-13.
6. Kim, E.J., Jeong, D.U., (1999). Transdermal scopolamine alters phasic REM activity in normal young adults. *Sleep*, **22(4)**, 515-520.

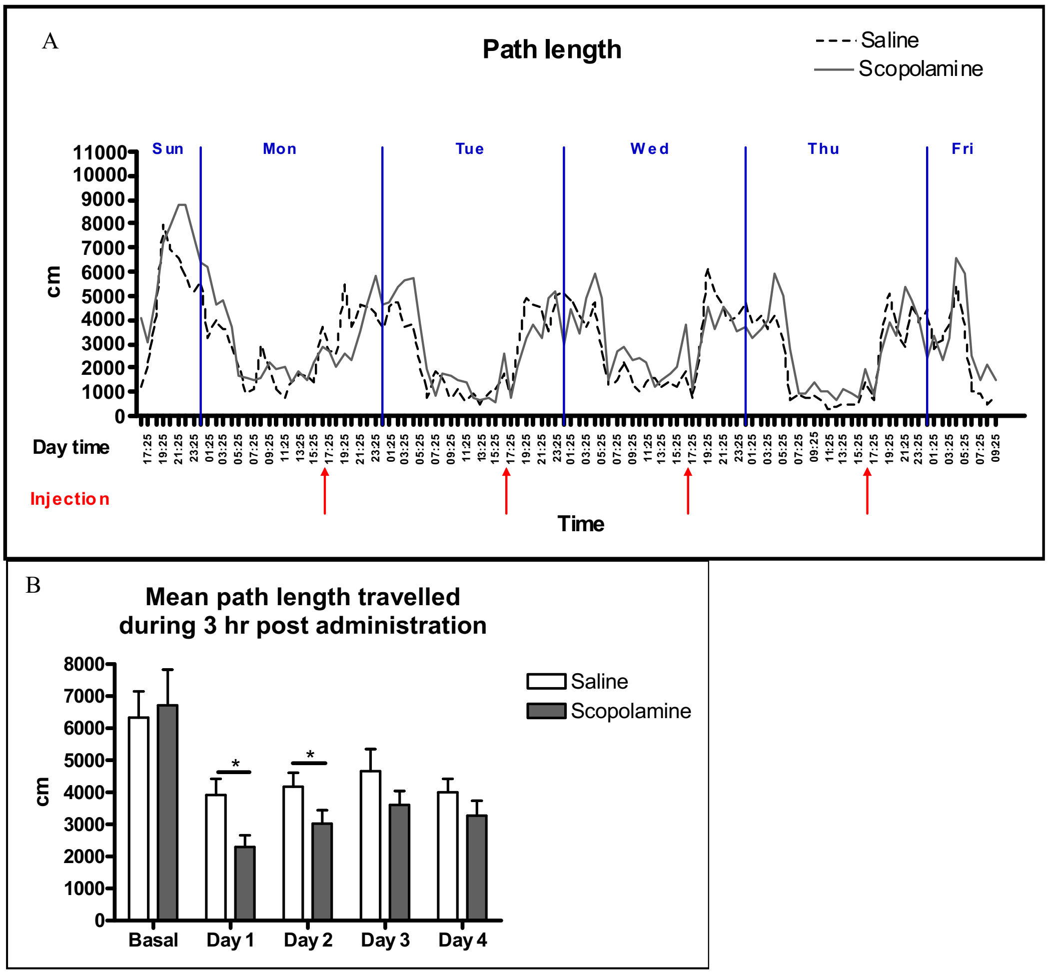


Figure 1. *Circadian rhythm recorded over four consecutive days in saline- and scopolamine-treated mice as changes in locomotor activity (a). Scopolamine did not change the normal circadian locomotor activity, but it decreased activity for 3 hr post-administration during the first two days of administration only (b). Asterisks represent significant difference at p<0.05.*

iBehave – application of supervised machine learning to behaviour analysis

J.A. Heward, P.A. Crook, T.C. Lukins, and J.D. Armstrong
School of Informatics, University of Edinburgh, Edinburgh, UK, douglas.armstrong@ed.ac.uk

Overview

Automatic recognition and classification of behaviour in laboratory animals is essential if behaviour research is to keep pace with other biological domains where the use of high throughput, data rich platforms are rapidly accelerating. In behaviour genetics studies and in the CNS drug discovery sector, behavioural throughput is already a rate limiting. There exists a number of approaches and in several cases implemented solutions for many common laboratory behaviours. In addition to fully automated systems there are partial solutions that maximise investigator productivity and enhance the accuracy and quantity of data returned.

The iBehave project is examining the application of supervised machine learning methods in this domain. The goal in simple terms is as follows. Given tracked data for an experiment, can we exploit the expertise of a human observer to train a computer algorithm to classify behaviour? We will present an overview of the current system along with demonstrations.

The software employs supervised machine learning methods to classify the behaviour of multiple complex objects in video footage. The principle advantage of our approach is that the system can be trained by the end-user to detect novel behaviours. This is achieved through training the system using selected exemplars identified by the end-user in pre existing footage. The software extracts coordinates and parameter values which uniquely distinguish the set of given exemplars. Thereafter, each behaviour that the system has been trained on can be automatically detected with an expression of confidence. The method can be to applied to footage containing both single and multiple articulated objects.

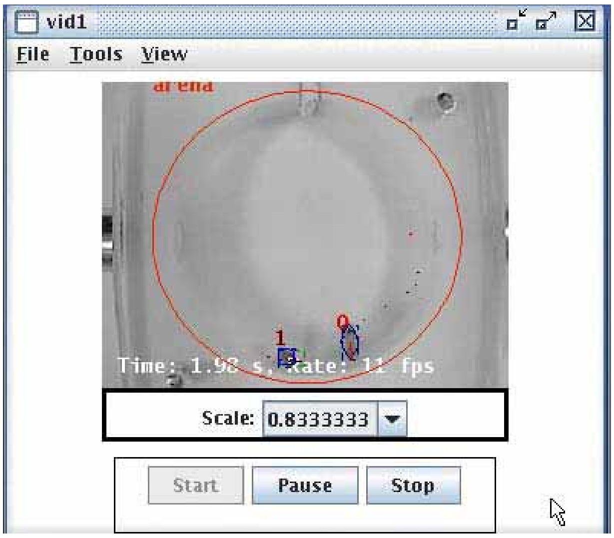


Figure 1 *Flytracker application following two adult fruit flies.*

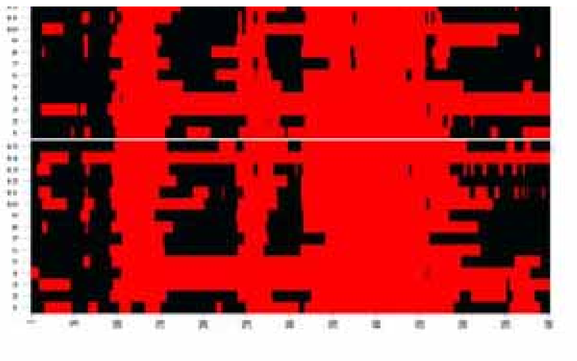


Figure 2*. 15 experts annotating the same 600 video frames. Red indicates the behaviour occurring, black no behaviour. The top 15 rows show the first presentation the lower 15 rows the same video rotated through 180 degrees and represented later. Inter-observer variation is much higher than intra-observer variation.*

Design

The methods require three conceptual components: Parameter Extraction, Data Classification, and Annotation. Our main activity is in the Data classification and Annotation components. However, detailed parameter extraction is essential to the success of the programme.

Parameter Extraction

In the small animal behaviour measurement scenario, the software receives raw data in the form of video footage. It is first necessary to reduce the data to a set of key parameters that model the behaviour of the objects under observation. In this case, the key parameters are coordinates describing the shape, relative movement and deformation of body parts over time. Extracting these parameters reliably, across a range of experimental conditions, is a highly non-trivial challenge.

Data Classification

The operator specifies the subject behaviours by clicking start and stop to define exemplars. The key parameters pertaining to these exemplars are stored in the software and converted into classification rules describing the global boundary conditions (of the subject behaviours).

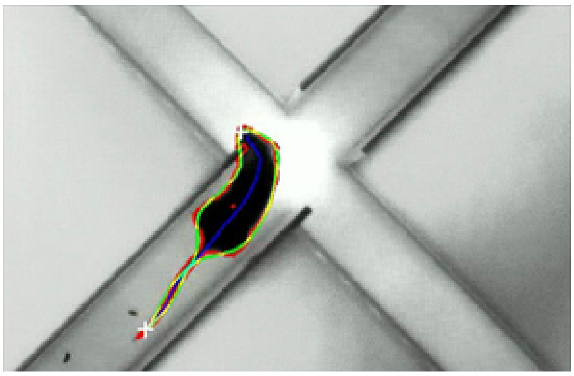


Figure 3. *Rodent behaviour requires additional parameters such as the body contours to be extracted.*

Annotation

The software automatically screens new footage against the existing classification rules to annotate episodes that appear to contain the subject behaviours. Measures for statistical confidence are returned with mathematical expressions describing the classification rules employed.

Case Studies

We created a test dataset using the courtship ritual of the fruit fly *Drosophila melanogaster* for a first case study. 15 experts from a number of labs across the world were recruited and all annotated the same selection of videos which included obvious as well as difficult examples of the behaviour in question (courtship index). We then tracked the x and y co-ordinates of each individual's centre of gravity as well as the angle of orientation (see figure 1), acceleration and velocity (both linear and angular) . From these additional relative parameters are calculated.

A total of 10 minutes of footage was shown to each expert comprising nine one minute clips, of 600 frames each, taken from nine different previously recorded five minute videos were used to create a ten minute video (with a 10 frame delay between end/start of consecutive clips). Only nine clips were needed because the tenth clip was an inverted version of the fourth. The clips were of wildtype male and female *Drosophila* in a cylindrical plexiglass 20mm diameter x 5mm deep observation chamber. We compared the annotation across the experts (figure 2) which showed broad, but not exact inter-observer agreement. The video that was shown twice (albeit transformed) confirmed that intra-observer variation was lower than inter-observer variation.

80% of the annotated data was sampled to train using a decision tree with 20% held aside for validation. This was repeated for 5 random samples giving a final accuracy of 84.88% (st dev 2.54). This represents an 'average expert'. Training accuracy using single experts returns a much higher accuracy (of the order 94%).

Discussion

The iBehave method has been successfully applied to fly courtship. In particular it has highlighted the difficulties in getting multiple experts to agree on a common behavioural interpretation. We have recently extended the system to look at rodent behaviour and added additional parameters for the more complex body shape (figure 3). This work is on-going and will be presented at the meeting.

Evaluation of space requirements of broiler chickens by analysis of their spatial distribution

S.A.F. Buijs[1,3], F.A.M. Tuyttens[1], J. Baert[2], J. Vangeyte[2], E. Van Poucke[1], and L.J. Keeling[3]
[1] *Animal husbandry and welfare, ILVO-Animal sciences, Melle, Belgium*
[2] *Agricultural engineering, ILVO-Technology and food sciences, Merelbeke, Belgium*
[3] *Department of Animal Environment and Health, SLU, Uppsala, Sweden*
stephanie.buijs@ilvo.vlaanderen.be

Introduction

Broiler producers across the world raise about 20 billion chicks each year and the stocking densities under which this production is achieved vary greatly between countries and husbandry systems. A major criticism of intensive husbandry in general is that it compromises animal welfare by providing insufficient space per animal [1]. However, for broiler chickens it has recently been claimed that environmental circumstances may be of greater importance [2]. The spatial requirements of broiler chickens have most often been studied by looking at the adverse physical effects of high stocking densities (for instance decreased walking ability, increased contact dermatitis and mortality) or by studying changes in the behavioural repertoire. Studies on spatial distribution are scarce although they offer opportunities to investigate the animal's spatial preference more directly. For example, if close proximity of pen mates is experienced as aversive by broiler chickens, they may position themselves further away from their conspecifics when given the opportunity to do so [3] (thus, at lower stocking densities). Furthermore, it is important to study the spatial distribution in association with behaviour, as the type of behaviour being displayed depends on this distance [4].

In this experiment the spatial distribution of broiler chickens over their home pen was studied at different densities and in relation to behaviour. The aim was to determine the density at which the distance to the nearest neighbour became equal to the distance expected purely by chance. It was assumed that this point would be informative of the spatial requirement of group housed broiler chickens. In addition, the number of birds in each quadrant of the pen was assessed to study whether this simplified measure of distribution yielded similar results.

Materials and methods

Broiler chickens were housed at 8 different densities: 8, 19, 29, 40, 45, 51, 61 and 72 birds / 3.3 m^2 pen corresponding with 6, 15, 23, 33, 35, 41, 47 and 56 kg live weight / m^2 at the end of the rearing period (day 39). Males and females were mixed at a ratio of 1:1. To avoid clustering around resources, water cups and feeders were distributed evenly over all sides of each pen and no lamp brooders were supplied. The ambient temperature was 31° C at 1 day of age and was lowered by 1° C every 3 days until a temperature of 21° C was reached. Lights were on for 21 hours per day. In each pen, 8 focal birds were colour marked to allow individual recognition. Each density treatment was replicated four times.

Pens were filmed from week 2 to 6 for one day per week, using an automated digital recording system. Each pen was recorded for 5 minutes at a time, 6 times per day (twice each morning, afternoon and evening). The first image of each recording was used to determine the spatial distribution of the chickens in the pen. Images were calibrated to minimize the amount of distortion due to the wide-angle lens of the camera and coordinates of each bird were noted using customized software built in Halcon 7.1 (see Figure 1). From these coordinates, the distance between each bird and its nearest neighbour was calculated. These distances were compared to those acquired from simulations in which chickens positioned themselves regardless of their distance to other chickens. These simulations were created by taking random samples from all pooled combinations of XY-coordinates for the particular density, week and replicate in which the experimentally observed distribution was determined. The number of simulations was determined by the number needed to achieve a stable average distance. In addition, the number of birds in each quadrant of the pen was acquired and compared to an equal distribution of animal over the quadrants.

Behaviour was studied in two ways: a continuous focal sampling was carried out to assess the frequency, total duration and bout length of the different types of behaviour of one focal bird per recording. In addition, a scan sampling of the first behaviour shown was carried out on all 8 focal birds in each recording. Behavioural data from the scan sampling was linked to that of the spatial distribution to study the interaction between behaviour and inter-individual distance. In both cases, two ethograms were used simultaneously; one that documented the animal's posture and locomotion (stand, sit, lie, walk, run and adjust) and one that documented their activities (eating, drinking, ground pecking, agonistic behaviour, ground scratching, preening, dustbathing, leg stretching, head flicking, comfort behaviours (wing flapping, body shaking and tail wagging), displacing a chicken from the feeder or drinker, being displaced from feeder or drinker and "other" (all behaviour that did not fall into the previous activity categories)). The total number of behavioural transitions was calculated from the combined frequency of the other types of behaviour.

The results acquired using these methods will be presented at the conference.

References

1. Scientific Committee on Animal Health and Animal Welfare (2000). The welfare of chickens kept for meat production (broilers).
2. Dawkins, M.S., Donnelly, C.A., Jones, T.A. (2004).Chicken welfare is influenced more by housing conditions than by stocking density. *Nature,* **427**, 342-344.
3. Keeling, L., (1995). Spacing behaviour and an ethological approach to assessing optimum space allocations for groups of laying hens. *Applied Animal Behaviour Science*, **44**, 171-186.
4. Keeling, L.J., (1994). Inter-bird distances and behavioural priorities in laying hens: the effect of spatial restriction. *Applied Animal Behaviour Science*, **39**, 131-140.

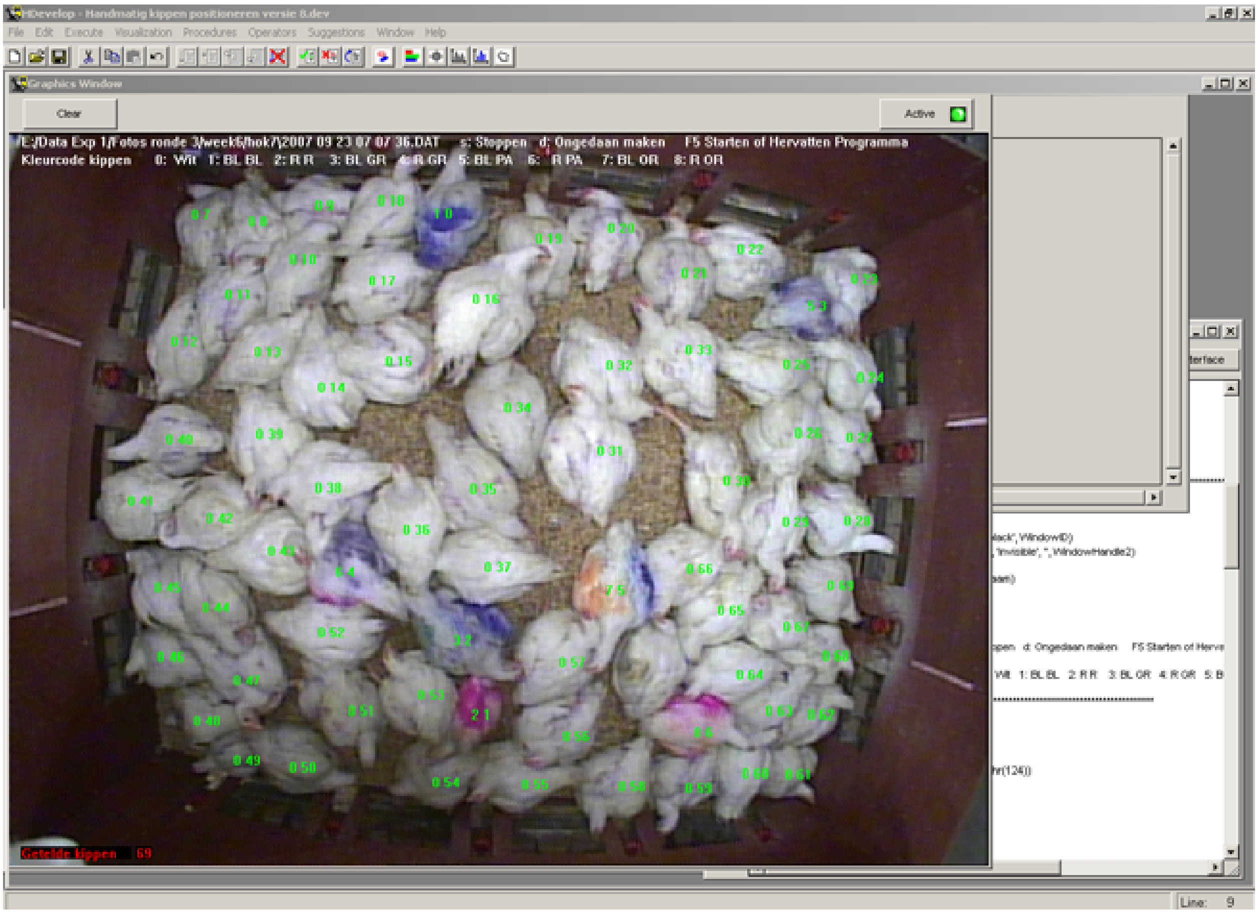


Figure 1. *The Halcon interface for the determination of the XY-coordinates of each broiler in the pen. The first figure of each number indicates the location on the bird where the coordinate was scored (i.e. centre of the animal).*

A comparative analysis of hierarchically organized grooming patterns in hamsters

K.V. Anokhin[1], T.V. Bogolepova[1], A.V. Surov[2], and I.Yu. Zarayskaya[1]
[1] *Department of Systemogenesis, P.K. Anokhin Institute of Normal Physiology, RAMS, Moscow, Russia*
[2]*A.N. Severtsov Institute of Ecology and Evolution, RAS, Moscow, Russia, k_anokhin@yahoo.com*

Introduction

Structural organization is a fundamental feature of animal behavior. It is usually a result of evolution or learning and is expressed in hierarchically organized successions of behavioral acts ("behavioral patterns"). Analysis of such patterns opens a possibility to interpret animal behavior, categorize it and to trace its history and evolution. However, the analysis of behavior structure and its transformations is a complicated task that requires development of new specific algorithms similar to those used for the study of morphological traits and DNA sequences.

Proposed approach

Our approach to behavior structure analysis is based on a functional systems theory (FST) [2-4]. According to the FST, behavioral continuum is separated into individual behavioral acts, each based on a discrete functional system [3]. Functional systems are dynamically self-organized aggregates of neural and peripheral elements that cooperate to achieve an adaptive result for a behaving organism [2]. Achievement of the result of a particular system terminates its activity and initiates self-organization processes for the next functional system [8]. Changes in dynamics of animal behavior at the moments of such critical transitions ("breakpoints") allowed us to develop motion analysis algorithms that isolate elementary behavioral units from a behavioral continuum [6]. Our next step, followed in the current work, was to reveal and reconstruct a natural structure of behavioral continuum as consisting of hierarchically organized patterns of behavioral acts ("systemic quanta" of behavior [9]).

To develop the new tools for detection and comparative analysis of such behavioral patterns we used highly organized rodent grooming behavior. The main features of grooming behavior are high level of movement stereotypy and fixed successions of discrete grooming acts, which can be defined as evolved actions sequences possessing syntactic rules [5]. Grooming patterns show intraspecies stability and interspecies variability in wide spectrum of taxons. These characteristics and possibility to induce this behavior under experimental conditions and to assess it quantitatively resulted in our selection of grooming as a model for analysis of behavior structure.

Experiments

We performed comparative structural analysis of grooming in several species of Palaeartctic hamsters (subfamily *Cricetinae*) belonging to five taxons with known phylogenic connections between them: *Phodopus campbelli, Mesocricetus auratus, Cricetulus griseus, Cricetus migratorius* and *Allocricetulus eversmanni*. Grooming behavior was induced in animals by water immersion. Each animal was videotaped during 15-min test session under room light in a glass chamber with two side mirrors. The video tracking of animal's movements was carried out using "Easy Track" software developed in our lab [1]. The subsequent segmentation of grooming continuum into single behavioral acts was performed automatically by "Segment Analyzer" software developed in our lab [6]. In total we detected 24 different grooming acts which were defined as combinations of movement category (licking, washing, biting etc.) and category of body region to which movement was confined ("hand licking", "forelimb biting", "snout washing", etc.). Acquired successions were analyzed by the Theme software to reveal the hierarchy-organized grooming T-patterns [7].

A comparative analysis of grooming patterns uncovered both differences and similarities between sequences of grooming acts in five species. Major difference was in the strategies of transitions from face and head grooming to other body regions. Another pronounced difference was in the probabilities of substitution of single movement category within a similar body region T-pattern ("biting" vs. "licking", "washing" vs. "fast washing" etc.). On the other hand, strong similarities were found in cheeks sacks grooming patterns in all studied palearctic hamsters. Next we performed a classification of behavioral patterns according to different order of the studied taxons: subfamily-specific, species-specific and individual grooming patterns. All observed species displayed similar simple subfamily-specific patterns during face and head grooming ("snout washing – fast snout washing", "snout - eyes washing - fast snout washing", "snout washing - head and ears washing", "snout washing – forepaws licking", etc.). Structural analysis also showed that different species of hamsters had numerous individualities in fixed actions programs of grooming behavior including amount of different pattern types, pattern length, pattern hierarchy levels, directions of body regions transitions and body regions attractors for particular patterns.

The results of our subfamily-specific and species-specific analysis suggest that evolution of grooming behavior in Palaeartctic hamsters consisted in: (a) appearance of novel grooming patterns, (b) expansion of grooming patterns along the rostra-caudal body axis, (c) elaboration of face and a head grooming behavior, (d) involvement of forepaws in hard-fixed actions programs and (e) substitution of forelimbs to forepaws use in functionally similar patterns (see Figure 1).

Conclusion and future work

We conclude that functional systems approach to analysis of behavior allows to reveal structure highlights fundamental features in the organization of grooming behavior. It allowed us to detect the "systemic quanta" of grooming behavior, reveal their grouping into fixed actions programs, describe these patterns quantitatively and apply to them a comparative phylogenetic analysis. In future we intend to use this approach to reconstruct the evolution of separate functional systems of grooming behavior based on comparison of subfamily-, species-specific and individual hierarchically organized patterns of behavioral acts in palearctic hamsters and knowledge of phylogenic connections between the examined taxons.

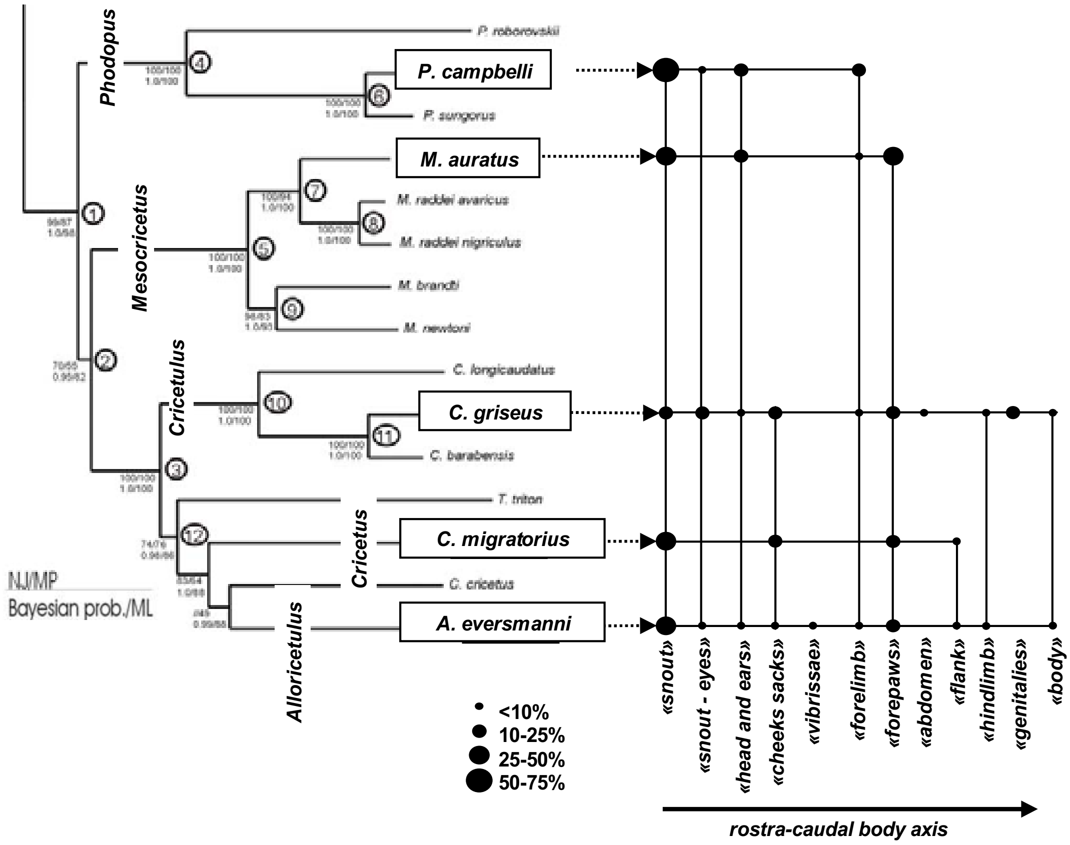


Figure 1. *A rostra-caudal expansion of species-specific grooming patterns topography (B) in concordance with molecular-genetic tree of phylogenic connections between the examined taxons (A).*

References

1. Anokhin K. and Mukhina T. (2005). Development of a neurocognitive interface for an animal video tracking system.; *Proceedings of Measuring Behavior 2005, 5th International Conference on Methods and Techniques in Behavioral Research*, Ed. by L. Noldus, F. Grieco, L. Loijens and P. Zimmerman. Wageningen: Noldus Information Technology.
2. Anokhin P.K. (1961). A new conception of the physiological architecture of conditioned reflex. In: *Brain Mechanisms and Learning*, Oxford: Blackwell, 189-229.
3. Anokhin P.K. (1968). Nodular mechanisms of functional system as self-regulatory apparatus. *Progress in Brain Research*, **22**, 230-251.
4. Anokhin P.K. (1974). Biology and Neurophysiology of Conditioned Reflex and its Role in Adaptive Behavior. Oxford: Pergamon Press.
5. Berridge K.C. (1990). Comparative fine structure of action: rules of form and sequence in the grooming patterns of six rodent species. *Behaviour*, **113 (1-2)**, 22-56.
6. Cherepov A.B., Mukhina T.V., Anokhin K.V. (2005). Automatic segmentation of mouse behavior during video tracking in home cages. *Proceedings of Measuring Behavior 2005, 5th International Conference on Methods and Techniques in Behavioral Research*. Ed. by L. Noldus, F. Grieco, L. Loijens and P. Zimmerman. Wageningen: Noldus Information Technology. p. 403.
7. Magnusson M.S. (1996). Hidden real-times patterns in intra- and interindividual behavior: Description and detection. *Eur. J. of Psychological Assessmen,* **12**, 112-123.
8. Shvyrkov V.B. (1986). Behavioral specialization of neurons and the systems-selective hypothesis of learning. In: *Human memory and cognitive capabilities*. Eds. Klix F., Hagendorf, North-Holland, 599-611.
9. Sudakov K.V. (1982). Functional systems theory in physiology. *Agressologie*, **23**, 167-176

Effect of lesions to the posterior parietal or medial prefrontal cortices on navigation based on distal or proximal orienting cues in the rat

J. Svoboda[1], P. Telenský[1,2], K. Blahna[1,2], J. Bureš[1], and A. Stuchlík[1]
[1]*Institute of Physiology, Academy of Sciences of the Czech Republic,Prague, Czech Republic, svobodaj@biomed.cas.cz*
[2]*Faculty of Science, Charles University, Prague, Czech Republic*

Despite the long-lasting research on rodent spatial memory, we are still unable to specify the particular roles of various cortical structures in a satisfactory manner. Here we examined the impact of lesions to the Posterior Parietal (PPC) and Medial Prefrontal (MPFC) cortices on two variants of rat place avoidance task, based on distal or proximal orienting cues.

Generally [1,2], the place avoidance apparatus consists of a smooth metallic rotatable circular arena (80 cm in diameter) enclosed with a 30 cm high transparent Plexiglas wall and elevated 1 m above the floor. The rat placed on the arena is wearing a light latex harness, to which an infrared light-emitting diode (LED) is attached between the rat's shoulders. Its position is tracked every 40 ms by overhead camera and recorded onto a computer track file, allowing subsequent reconstruction of the track with an off-line analysis program (TrackAnalysis, Biosignal Group Corp., USA). Second LED is attached to the outer part of arena circumference serving as a referencing point when monitoring position of the animal within rotating arena frame. Thus, the system allows to store and analyze animals' tracks both in arena and room frame coordinates. Animals are trained in daily 20 min sessions to avoid entering a directly imperceptible 60-degrees sector of the arena, defined either in room or arena reference frame (see below). Whenever a rat enters the shock sector for more than 0.5 s, mild electric shocks (50 Hz, 0.5 s) are delivered at intervals of 1.5 s unless the rat leaves the shock sector for at least 0.5 s. The shocks are delivered through a thin subcutaneous low-impedance nichrome wire implant on the back of the rat standing on the grounded floor. The appropriate shock current (ranging between 0.2-0.7 mA) can be individualized for each rat to elicit a rapid escape reaction but to prevent freezing.

Here we present two variants of place avoidance task, differing in the spatial properties of the to-be-avoided sector. In the arena frame (AF) task, the sector is anchored to slowly rotating (1rpm clockwise) arena surface. Thus, rats have to remember position of the punished sector with respect to cues present on the arena surface (urine marks, etc.).To motivate rats to actively and homogenously explore the arena, they are run on a restricted diet schedule and left to search for peeled barley grains scattered at 10s intervals randomly over the arena from centrally placed overhead feeder. Since the arena is slowly rotating rats must abandon the use extramaze cues. On the contrary, in room frame (RF) task, the punished sector is anchored to reference frame of the room and therefore does not move with rotating arena. In this case, rats are required to use distal (extramaze) navigational cues such as doors, shelves, etc. and to abandon navigation based on proximal (intramaze) cues.

Following measurements were evaluated in our experiments: number of entrances into punished region (expressed learning rate), and total path during the session (expressed the overall locomotor activity). To eliminate possible confounding effect of variable locomotor activity during the AF task, number of entrances divided by path length was added. In RF task we also measured maximum time the animal was able to avoid entrance to the shock sector within a session. This variable is applicable only in RF design, where continuously rotating arena inevitably brings an animal sitting immobile to the punished region. The immobile animal sitting outside punished region in AF task is, however, undistinguishable from completely avoiding rat on the basis of max t evaluation, thus making max t uninformative. All variables were statistically analyzed using two-way ANOVA (lesion x sessions) design with days as repeated measures.

Rats with thermocoagulation lesion of the PPC (n=10) did not differ from matched group of sham-operated controls in their performance (number of entrances or overall locomotion) in AF nor in RF variant of the task. Preliminary experiments with MPFC lesioned rats led to the same results. These data suggest that PPC and MPFC are not crucial for navigation based neither on proximal nor distal orienting cues, at least at the level of precision given by the accuracy demands of the used behavioral paradigm.

All experiments were in accord with the directive of the European Communities Council No. 86/609/EEC. Supported by GACR grants 309/07/0341, 309/06/1231 and 206/05/H012, MSMT CR 1M0517.

References

1. Cimadevilla, J.M., Kaminsky, Y., Fenton, A., Bures, J. (2000). Passive and active place avoidance as a tool of spatial memory research in rats.*J Neurosci Methods*, **102**, 155-64.
2. Cimadevilla, J.M., Wesierska ,M., Fenton, A.A., Bures, J. Inactivating one hippocampus impairs avoidance of a stable room-defined place during dissociation of arena cues from room cues by rotation of the arena.*Proc. Nat. Acad. Sci. USA*, **98**, 3531-6.

An efficient low cost system for fish swimming behavior analysis in ecotoxicology

P.S.M. Carvalho[1], M. Mostaert[1], P. Pessoa[1], A. Pedrosa[1], J.R.M.B. Correia[1], A.R. Ribeiro[1], and R. Bose[2]
1Department of Zoology, Universidade Federal de Pernambuco, Recife, Brazil, pcarvalho@ufpe.br
2Instituto de Engenharia Biomedica, Universidade Federal de Santa Catarina, Florianopolis, Brazil

Introduction

Fish are important model organisms in Ecotoxicology, especially during early developmental stages, when they are more sensitive to chemicals that contaminate fresh and marine waters [1]. Behavioral tests using these stages can be directly related to fundamental processes such as growth, survival, and recruitment to the adult population, therefore providing much needed links of suborganismal effects to population level effects of contaminants [2; 3]. The use of replication in gathering behavioral data in Ecotoxicology is difficult because of the need for efficient systems to record and analyze video images. Fish swimming behavior is essential for proper detection of prey food as well as for adequate behavioral avoidance of predators, and represents an important topic to be analysed in detail in Ecotoxicology [4]. Different automated systems based on either analog or digital video recordings of fish locomotor behavior, followed by manual or automated analysis based on commercial or academic software have been proposed and used in the literature [5; 6; 7]. Many of these systems were not specifically interested in applications related to fish early development, when their small size makes the obtention of good quantitative information on their swimming behavior a more challenging task. In addition, commercially available software such as Noldus Ethovision, although efficient and reliable, is beyond the budget for many laboratories. Our objective in this study was to establish a digital video acquisition system and a software based analysis of swimming speed and distance traveled by small fishes that can provide reliable data at a reasonable cost for lower budget behavioral laboratories.

Methods

The video system used was based on the simultaneous capture of images from 16 black and white CFTV cameras (MTC, model KBC-602) with 6-60mm auto-iris manual zoom lenses connected to a 240fps video capture card model Geovision GV-800 (15fps per camera), installed on a desktop computer. The imaging system allows proper monitoring from a superior view of testing arenas ranging from 2 cm (length) x 1.5 cm (width), and up to 14 cm x 10.5 cm (used here) and larger, if necessary. Small fish *Poecilia reticulata* in the range of 1 cm total length were used in the experiments. A total of 60 fish were exposed to 5 concentrations of the organophosphorus insecticide DDVP, ranging from zero in controls, and 10, 100, 500, 1000 µg.L-1 during 96 hours. At 96 hours, 12 fish from each treatment had their swimming behavior in the test arenas individually recorded by the camera system during 5 min. Digital video images were analyzed by the software SpyNeuroTracking, which uses several attributes of the animal image in the video. Mass center position, angular position, and area are extracted from these records by the developed software that processes individually each frame of the video file. Post-processing of the raw data is carried out by the software to calculate behavioral attributes of interest within defined integration intervals for the test duration, including a tracking description map, velocity, distance and orientation of locomotion, time spent in predetermined areas, among others [8]. In order to validate the quantitative measurements by the software, 46 videos of 20s were selected for a comparison between the linear distance quantified by the Spyneurotracking software based on digital image analysis, and the linear distance graphically measured based on the tracking map generated by the software, which was checked to ensure the map was accurate in representing the fish path during selected videos of 20s.

Results and discussion

During the testing period, average linear distance swam varied from 8.2 cm/10s in controls to 13.2 cm/10s at the 500 µg.L-1 DDVP concentration. Swimming speeds varied from 0.8 cm/s in controls to 1.3 cm/s at the 500 µg.L-1 DDVP concentration (Figure 1). Although there is a tendency of hyperactivity in terms of an increase in both linear distance swam and swimming speeds of exposed fish, no statistically significant differences among the treatment groups were found (Anova, $p>0.05$).

Graphically measured linear distances swam (LDS-graphic), operationally considered the true linear pathway followed by the fish, were overestimated by 22% by the software quantified linear distances (LDS-software) for fish with low activity (swimming less than 3 cm linear distance during 20s). The same comparison based on fish that had medium or high activity (swimming more than 3 cm linear distance during 20s) indicated an agreement between both measurements within 2% (Table 1). In spite of the overestimation of distances swam for low activity fish, the data generated by the software analysis is reasonably accurate, validating the use of this system in modeling fish behavior in Ecotoxicology.

References

1. Weis, J.S., Weis, P. (1989). Effects of Environmental Pollutants on Early Fish Development. *Reviews in Aquatic Sciences*, **1(1)**, 45-73.
2. Little, E., Fairchild, J.F., DeLonay, A.J. (1993). Behavioral methods for assessing impacts of contaminants on early life stage fishes. *American Fisheries Society Symposium,* **14**, 67-76.
3. Scott G.R., Sloman K.A. (2004). The effects of environmental pollutants on complex fish behaviour: integrating behavioural and physiological indicators of toxicity. *Aquatic Toxicology,* **68(4)**, 369-92.
4. Little, E.E., Finger, S.E., (1990). Swimming Behavior as an Indicator of Sublethal Toxicity in Fish. Environ. *Toxicol. Chem.*, **9**, 13-19.
5. Vogl C, Grillitsch B, Wytek R, Spieser OH, Scholz W. (1999). Qualification of spontaneous unidirected locomotor behavior of fish for sublethal toxicity testing. Part I. Variability of measurement parameters under general test conditions.**18(12)**, 2736-42.
6. Kane, A.S., Salierno, J.D., Gipson, G.T., Molteno, T.C., Hunter, C. (2004). A video-based movement analysis system to quantify behavioral stress responses of fish. *Water Res.*, **38(18)**, 3993-4001.
7. Blaser R, Gerlai R. (2006). Behavioral phenotyping in zebrafish: Comparison of three behavioral quantification methods. *Behavior Research Methods*, **38**, 456-69.
8. Bose, R. (2005). A System for Tracking and Movement Analysis of Laboratory Animals Based on Image Processing. *Revista Brasileira De Engenharia Biomédica*, **21(1)**, 61-72.

Table 1. *Comparison of software and graphically based calculation of linear distance swam (LDS) by individually tested fishes*

Activity level	LDS-software (cm)	LDS-graphic (cm)	Ratio software/graphic
low	n = 12		
0-3cm/20s	1.65	1.35	122%
medium	n = 14		
3-10cm/20s	9.42	9.61	98%
high	n = 20		

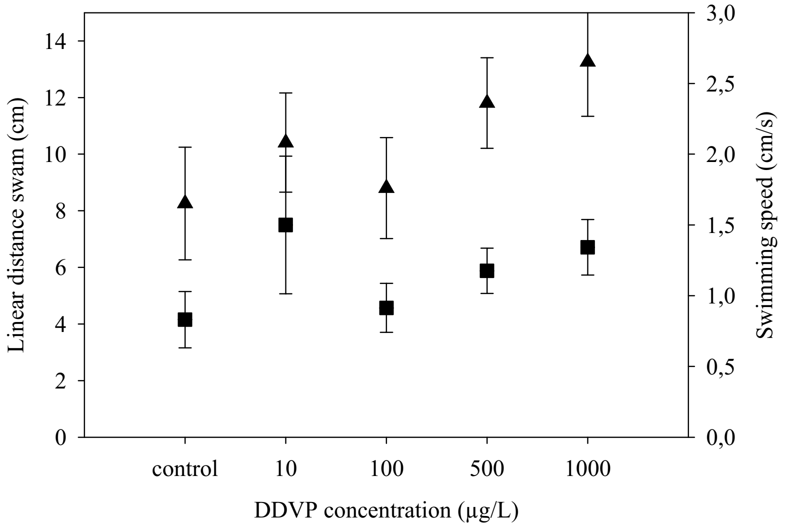


Figure 1. *Linear distance swam (triangles) and swimming speed (squares) of Poecilia reticulata exposed to DDVP based on video acquisition system and Spyneurotracking software analysis*

The effects of prenatal morphine exposure on spatial learning

A. Alijarahi

Departement of Physiology, Qazvin Azad University, Qazvin, Iran, A.Alijarahi@yahoo.com

Background

Drug abuse during pregnancy is a growing problem in all developed countries of the world. Maternal drug abuse affects the developing system and its long-term effects can persist till adulthood so it can decreases the rate of their maturation..Thus the present study was designed to determine whether the exposure to the morphine during gestation permanently alter spatial-learning.

Objective

To determine the effects of prenatal morphine exposure on spatial learning

Materials and Methods

18 Pregnant rats were divided to morphine, saline and control groups. Morphine or saline were administrated (S.C) to female rats twice a day (08h and 20h) on gestational days 11-18, (5 mg/kg morphine for 3 days and 10mg/kg for 5 days). pups (P90, n=6) were trained in an 8-arm radial maze apparatus.

Findings

The results indicate that prenatal morphine exposure reduces the time needed to learn this trials, but they needed more time to complete regular trials. **Conclusion**: Prenatal morphine exposure impairs normal spatial learning.

Keywords

Morphine, Prenatal-spatial learning-rat.

The rat formalin test: Can it predict neuropathic pain treatments?

A. Ellis, N. Benson, I. Machin, and L. Corradini
Pfizer Global Research and Development, Sandwich, UK, amanda.ellis2@pfizer.com

Widely used, the formalin test is a tonic model of continuous pain resulting from formalin-induced tissue injury. It is a useful model, particularly for the screening of novel compounds, since it encompasses inflammatory, neurogenic, and central mechanisms of nociception [5,6]. Reports can vary widely with regard to concentration of formalin (itself a 37% dilute solution of formaldehyde) as well as the recording and characterisation of resultant pain behaviours in rats [4,5,6]. Therefore the effect of 1.25, 2.5, and 5.0% formalin on nociceptive behaviours was evaluated by measuring two commonly reported pain-like behaviours – flinching and licking/biting [1,2,3], associated with formalin intraplantar injection. Three clinically effective neuropathic pain treatments were tested (Gabapentin, Duloxetine, and Mexiletine) [1,2,3] to determine if these could reduce formalin-induced pain-like behaviours. An electronic recording method (Noldus Observer 5.0) was explored as an alternative to manual methods using observation and stop clocks. This allowed simultaneous recording of licking/biting and flinching behaviours.

Methods

All procedures were carried out under the Animals (Scientific Procedures) Act, 1986. Male Sprague Dawley rats (200-300g, Charles River, UK) were housed in groups of 4 under a 12 hour light/dark cycle with food and water *ad libitum*. All experiments were carried out by an investigator blind to drug treatments.

Formalin test

Rats were placed singly in perspex boxes (30 x 30 x 30cm, with a mirror placed on the back panel to aid observation) for approximately 15 minutes before the start of the test for habituation. Animals were pre-treated with standard analgesics at different time points according to existing pharmacokinetic/pharmacodynamic (PKPD) data (data not shown). Rats then received intraplantar injection into the right hindpaw of a 50ul solution of either 1.25, 2.5, or 5.0% formalin solution (saline vehicle) using a 29G needle, and placed immediately back in the boxes. Pain-like behaviours (licking/biting and flinching) were recorded in 5 minute time bins using the Pocket Observer (Noldus) for 45 minutes. Data were also considered in either the Early Phase (0-10mins) or the Late Phase (11-45mins). Animals were euthanased immediately at the end of the study. All experiments were recorded using a video camera as back up information.

Noldus Observer PDA

The PDA was pre-programmed to record two rats simultaneously for each investigator, and the behaviours of interest – licking/biting, flinching, and “rest” of the ipsilateral paw. To record a behaviour the appropriate subject and then behaviour, was selected on the screen as the behaviour began (ie Rat 1, Lick). At the end of the experiment the data were uploaded to a PC which was programmed to calculate the durations of each occurrence of recorded behaviour within user-defined intervals (in this case 5 minute time bins). This information was then exported to a Microsoft Excel spreadsheet for final analysis of this raw data.

PKPD and Statistical Analysis

Each treatment group was compared at each time point, as well as the Early and Late phases, to vehicle treated groups using one-way ANOVA and post-hoc student’s t-test. Results were expressed as mean ± S.E.M.

PKPD data analyses were carried using NONMEM V5 and S-Plus V 6.1 (Insightful corp, US).

Results

Unilateral intraplantar injection of formalin in the rat generates a biphasic pain- like behaviour characterised by licking/biting/ and flinching. . A comparison with the effect of higher (5%) and lower (1.25%) doses of formalin, suggested 2.5% solution generates a robust effect and acceptable levels of variability. Therefore this dose was chosen for the pharmacological characterisation of the model. Data obtained using stop clocks were found to be consistent with those obtained using the Pocket Observer, and not significantly different at any time point (data not shown). Therefore further data capture was performed using the Pocket Observer allowing simultaneous comparison of both licking/biting and flinching end-points.

Three clinically effective neuropathic pain treatments (Gabapentin 100mg/kg p.o., Duloxetine 1, 3, 10mg/kg i.p., and Mexiletine 10 and 30mg/kg p.o.) [1,2.3] were tested. Although no significant effect was seen on flinching behaviour, all three compounds reduced the duration of licking/biting behaviour in the second phase of the formalin test. Gabapentin, Duloxetine, and Mexiletine significantly ($p=<0.05$) reduced pain-like behaviour at 100 mg/kg, 10 mg/kg and 30 mg/kg respectively. Further PK-PD analysis is ongoing in order to establish whether this model can be used to predict dose translation from rat to human. The results of the rat model PKPD analyses will be presented in the context of clinically used exposures.

In conclusion the rat formalin test, using a 2.5% solution with licking/biting as the measured end-point, represents a suitable model to explore new targets and establish their potential as treatments in neuropathic pain. This data supports similar findings previously demonstrated by Vissers et al. showing pharmacological correlation between the formalin test and neuropathic pain behaviours [7]. Ongoing PKPD analysis will help to establish whether this model can also be a suitable tool to help dose prediction from rat to man.

Finally, the Observer represents a useful tool for electronic data storage which is more temporally detailed than traditional methods. Advantages of this method included the electronic storage of raw data (detailing timing of recorded events), as well as accuracy and objectivity of recordings. Additionally it allows the observation of multiple behaviours in multiple subjects, using just one PDA. Although convenient, further development of the Observer software would be useful to speed up the process of transferring data into a suitable format for graph and statistical analysis.

References

1. Dixit, R., Bhargava, V.K., Kaur, N. (1999). Antinociceptive effect of gabapentin in the formalin test. *Methods and Findings in Experimental and Clinical Pharmacology*, **21**, 481-482.
2. Iyengar, S., Webster, A.A., Hemricke-Luecke, S.K., Yu Xu, J., Simmons, R.M.A. (2004). Efficacy of Duloxetine, a Potent and Balanced Serotonin-Norepinephrine Reuptake Inhibitor in Persistent Pain Models in Rats. *The Journal of Pharmacology and Experimental Therapeutics*, **311**, 576-584.
3. Jett, M.F., McGuirk, J., Waligora, D., Hunter, J.C. (1997). The effects of mexiletine, desipramine, and fluoxetine in rat models involving central sensitisation. *Pain*, **69**, 161-169.
4. Kaneko, M., Hammond, D.L. (1997). Role of Spinal γ-Aminobutyric $Acid_A$ Receptors in Formalin-Induced Nociception in the Rat. *The Journal of Pharmacology and Experimental Therapeutics*, **282**, 928-938.
5. Lee, I.O., Kong, M.H., Kim, N.S., Choi, Y.S., Lim, S.H., Lee, M.K. (2000). Effect of different concentrations and volumes of formalin on pain response in rats. *Acta Anaesthesiologica Sinica*, **38**, 59-64.
6. Tjølsen, A., Berge, O.G., Hunskaar, S., Rosland, J.H., Hole, K. (1992). The formalin test: an evaluation of the method. *Pain*, **51**, 5-17.
7. Vissers, K.C.P., Geenen, F., Biermans, R., Meert, T.F. (2006). Pharmacological correlation between the formalin test and the neuropathic pain behaviour in different species with chronic constriction injury. *Pharmacology, Biochemistry and Behaviour*, **84**, 479-486.

An original method to interpret neurobehavioral data generated by the Irwin test in the mouse

P. De Ron, A. Delaunois, E. Hanon, Y. Lamberty, and M. Guyaux

UCB Pharma, Department of Non Clinical Development, Brainel'Alleud, Belgium, pierrette.deron@ucb-group.com

The Irwin test is a systematic observational procedure for assessing and scoring the effects of drugs on the behavioral and physiological state of rodents. This method described by Irwin (1962, 1968) [1] is part of the safety pharmacology core battery recommended by the ICH (International Committee for Harmonization) to detect potential adverse effects of candidates on the central nervous system (CNS) before human testing. When applied at an early stage of drug development, this test is particularly suitable to: a) screen and select compounds against unwanted CNS effects, b) understand the mechanisms underlying these effects, c) help improving the structure-activity relationship, and d) possibly, reveal novel therapeutical effects. By using an appropriate dose range for each test molecule, it is possible to obtain information on its pharmacological profile, on the intensity and the duration of its effects, and on the specificity of these effects ("on- or off-target" effects). At UCB Pharma, we refined the original Irwin test to better fit to the 3R's rules, by reducing the number of animals per dose, and by avoiding too high doses which could induce severe signs or lethality. NMRI mice of around 20 g receive the test compound by intraperitoneal route (n=3 per dose) and are observed at 0, 5, 15, 30, 60 and 120 minutes after drug administration. At each observation time, 53 parameters are scored using a rigorous standardized procedure based on that described by Irwin, 1968. These parameters are distributed as followed: 15 items for the behavioral profile, 21 items for the neurological profile and 17 items for the autonomic profile. A minimal active dose can be determined for each parameter. The originality of our interpretation of the data, collected as scores, is that some of them (36 items) are categorized into 5 fundamental activities: Central Activity (CA), Central Reactivity (CR), Neurovegetative Reflexes (NR), Neuromotor Tonus (NT), and Autonomic System (AS). A large set of reference substances, reflecting a large panorama of different pharmacological activities, were screened in this Irwin test. After a close scrutiny of the affected behavioural parameters with these reference substances, we developed an algorithm for the interpretation of their effects (increase or decrease) on the 5 "fundamental" systems (CA, CR, NR, NT and AS). Each new substance tested with unknown target, affecting some of the 53 items, can be submitted to the algorithm, in order to identify its physiological and pharmacological activities, and compare its profile with that of products already known.

References

1. Irwin, S. (1968). Comprehensive observational assessment: Ia. A systematic, quantitative procedure for assessing the behavioral and physiological state of the mouse. *Psychopharmacologia (Berl.),* **13,** 222-257.

Recording and tracking of locomotion and clustering behaviour in young honeybees (*Apis mellifera*)

M. Szopek, G. Radspieler, T. Schmickl, R. Thenius, and K. Crailsheim
Department of Zoology, Karl-Franzens University, Graz, Austria, grrad@hotmail.com

Introduction

Comprehensive studies using a one-dimensional temperature gradient revealed, that freshly emerged honeybees display a preference for temperatures similar to the one measured in the brood nest. For *A. mellifera* this temperature was found to be at approximately 36 °C [1]. Subsequent experiments in temperature organs, closed glass cylinders in which bees are exposed to a rather one-dimensional gradient, showed that single bees are able to find the optimum temperature in a steep gradient.

In contrast to that, recent studies at our department, using a two-dimensional gradient, showed that single bees were not able to find the optimum neither in flat nor in intermediate gradients, while sufficiently large groups of bees preferably form clusters in areas at the optimum temperature in intermediate gradients. We interpret this ability as an effect of swarm intelligence and the aim of our present study is to determine the basic mechanisms of this complex behaviour.

Material and Method

During the early stage of development of our experiment we encountered several challenges: We had to stabilize the gradient within sufficiently narrow confines by keeping a constant room temperature. In order to improve the camera's image quality we needed to illuminate the scene without shadows. Insufficient resolution of the camera posed a major problem for the tracking algorithm because the used tracking program was not able to distinguish between two or more bees which are too close. At present our experimental setup consists of the following components:

We observe the bees in a circular arena with a diameter of 60 cm, surrounded by a plastic wall with a height of 9.5 cm. To make sure the bees remain in the arena, we coated the wall with Teflon-spray. Covering the arena is not necessary, because freshly emerged bees are not yet able to fly.

The arena is situated in a small room on the top of a table. We are able to control the ambient temperature within narrow constraints to generate various temperature gradients in combination with one or two heating lamps which are located at different positions above the arena. We use ceramic heaters which are normally used for terrariums. To survey the temperature gradient in the arena we use an array of 64 highly sensitive temperature sensors which slightly protrude from the ground. We use self-designed electronics and software to collect the sensor data which are recorded and used as a feedback for automatically stabilizing the temperature gradient by controlling the power of the ceramic heaters. The temperature data are also stored for later evaluation. To provide an experimental environment that is close to natural conditions in the hive, we refrain from using visible light and use infrared light to illuminate the scene, because light with wavelengths beyond 660 nm is invisible for bees [2]. The 6 lamps for illuminating the recording area consist of halogen light lamps with infrared filters mounted in front of the lamp in a way that no visible light can shine through and they are distributed regularly around the arena.

For recording the bees' behaviour we use an infrared-sensitive surveillance camera which is fixed 70 cm above the centre of the arena. We record the videos digitally on a HD-recorder and extract one frame per second and process each frame with a self-written MatLab program, which computes the position of every single bee in the arena over time. Each bee is identified by computing the difference between the frame containing the bee and a reference frame of the empty arena. Possible noise is excluded by introducing a threshold for bee detection. The program produces an Excel file containing the positions of the bees at every time step and the average temperature of every sensor during recording time. We use Visual Basic to further compute the local temperature of the bees' positions, the distance to the optimum temperature (i.e. 36 °C), the distance each bee covers per time step, the angular deviation relative to the previous position and the number of bees in every potential cluster. The whole setup enables us to automatically record and evaluate a large number of samples of the bees' behaviour.

Further studies

In order to achieve even better results, we will make use of the currently evolving recording technology and employ higher resolution cameras. Further improvements to the tracking algorithm will allow us to identify cluster sizes with less or no error, and our current experimental setup will be extended to fit the needs of a wider array of research interests.

The parameters describing the bee's behaviour, which are a results of our experiments, will be included in the studies and experiments of the A*rtificial Life Lab* in our department where the derived algorithms will be implemented into Jasmin III robots [3].

References

1. H. Heran (1952). Untersuchungen über den Temperatursinn der Honigbiene (*Apis mellifica*) unter besonderer Berücksichtigung der Wahrnehmung von Strahlungswärme. *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, **34**, 179-207.
2. K. von Frisch (1965). Tanzsprache und Orientierung der Bienen. *Springer Verlag*, 487.
3. www.swarmrobot.org

The use of integrated behavioral station in chronic behavioral studies

Boris Sakic
McMaster University, Hamilton, Canada, sakic@mcmaster.ca

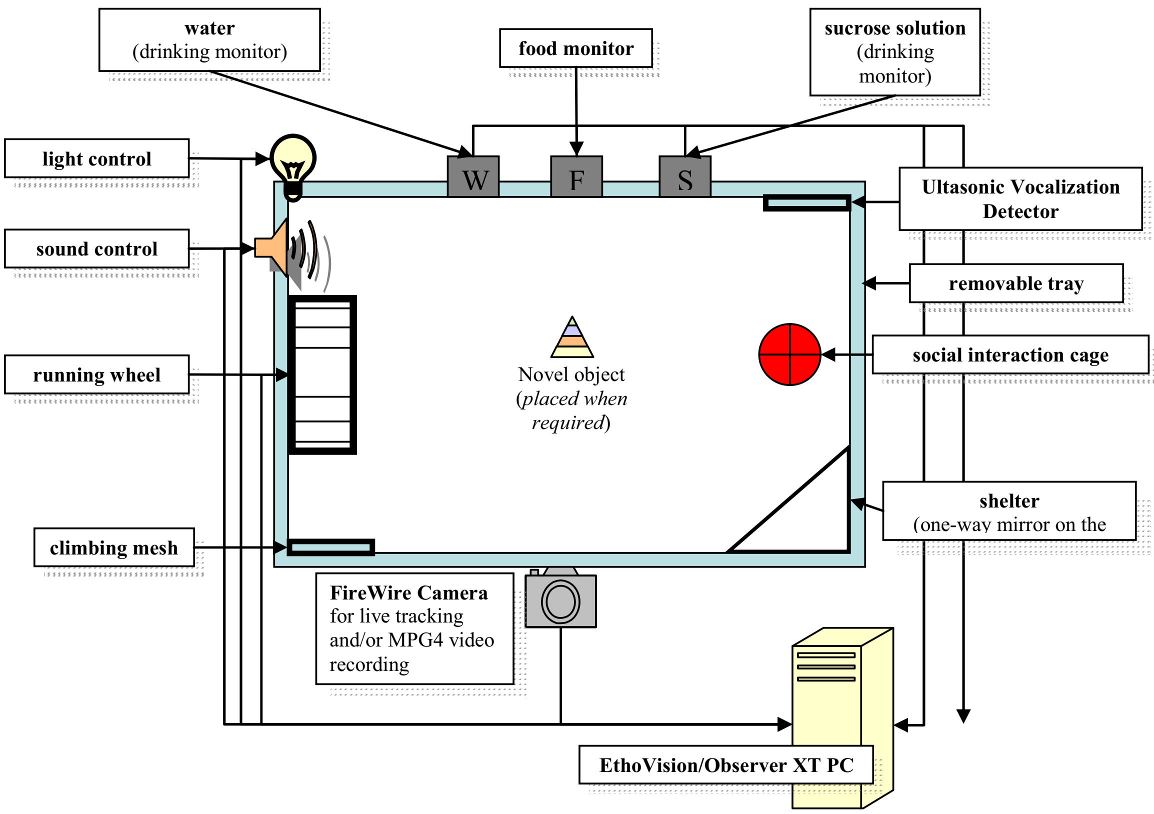
Human and animal neurodegenerative diseases are often accompanied by constellations of neurologic and behavioral deficits. In a murine model of neuropsychiatric lupus we observe age-dependent deficits in food/water intake, spontaneous locomotor activity, motivated behavior, emotional reactivity, and learning/memory performance. Anecdotal observations include seizures and self-injurious behavior [1]. However, given relatively short observations in behavioral tasks, no information could be systematically collected with respect to resting, sleep, food/water intake, or epileptic episodes. These important measures can be confounding factors when assessing performance in activity-demanding tasks, such as open-field, plus-maze, water maze, and forced swim tests [2]. This incompleteness in assessing behavioral profile of lupus-prone mice led to a necessity for a continuous, home-based monitoring of singly-housed mice during prolonged progress of systemic autoimmune disease. A custom-made integrated behavioral station (INBEST) was designed to meet this demand by concurrent, 24/7 measurements of multiple behavioral outputs in enriched home environment (see *Figure 1*). Dependent variables include measurements of food/water intake, responsiveness to palatable stimulation [3, 4], spontaneous ambulatory activity, climbing, voluntary running, anxiety-related behaviors, social interactions, sleep, and vocalization. Attached visual and auditory stimuli can be also used for conditioning and learning paradigms. The advantage of INBEST in comparison to standard behavioral testing is elimination of confounding effects induced by transportation stress and continuous, automated collection of measures reflective of nocturnal activity, exploration, anxiety-related and depressive-like behaviors. Third-party components and infrared MPEG4 technology integrated with EthoVision XT R2 and Observer XT are expected to yield wealth of information which will better account for onset, kinetics, and severity of behavioral changes, as well as important relationships among various behavioral deficits. The usage of INBEST can be expanded to other models of chronic CNS disorders, as well as to developmental, neuroethological, and chronic pharmacological studies.

This work was supported by the grant from Canadian Institutes of Health Research.

References

1. Chun S, McEvilly R, Foster JA, Sakic B. (2007) Proclivity to self-injurious behavior in MRL-lpr mice: implications for autoimmunity-induced damage in the dopaminergic system. *Mol Psychiatry* Sep 4.
2. Sakic B, Szechtman H, Talangbayan H, Denburg SD, Carbotte RM, Denburg JA. (1994) Disturbed emotionality in autoimmune MRL-lpr mice. *Physiol Behav*, **56(3)**, 609-17.
3. Sakic B, Denburg JA, Denburg SD, Szechtman H. (1996) Blunted sensitivity to sucrose in autoimmune MRL-lpr mice: a curve-shift study. *Brain Res Bull*, **41(5)**, 305-11.
4. Monleon S, D'Aquila P, Parra A, Simon VM, Brain PF, Willner P. (1995) Attenuation of sucrose consumption in mice by chronic mild stress and its restoration by imipramine. *Psychopharmacology (Berl)*, **117**, 453-7.

Figure 1. *Integrated Behavioral Station (InBeSt)*



Effects of environmental enrichments and olfactory stimuli on sleep pattern: an application of integrating information on behaviour and brain electrical activities

Chiara Rostello[1] and Elena Moscardo[2]

[1]*Animal Model Investigator Support Group, Laboratory Animal Science and* [2]*Safety Pharmacology Group, Dept. Safety Assessment GlaxoSmithKline R&D Centre, Verona, Italy*

chiara.b.rostello@gsk.com and elena.2.moscardo@gsk.com

Following differential experiences provided by enriched environmental conditions, several studies have shown significant changes in laboratory animals at morphological and behavioural levels [1, 2], but only few and not recent works evaluated effects of these conditions on sleep pattern [3, 4]. Besides, some data suggested the correlations between cortex, hyppocampal electrical activities and olfaction [5].

The aim of this study was to identify possible modifications of the behavioural pattern and/or of the concurrent recorded electroencephalographic traces in rats as a consequence of different environmental conditions (i.e., enriched or not enriched) and in the presence of olfactory stimuli applying the concept of integrating acquisition and analysis of physiological parameters with the behaviour of laboratory animals. In this research we used an integrated system for the simultaneous and synchronous recording of the behaviour by video (by Noldus Information Technology) and physiological EEG and EMG signals by telemetry (by Data Science International), which was installed and validated in our laboratory.

Four male CD™ rats, approximately 7-8 weeks old, were selected for the surgical instrumentation with a telemetric transmitter (type TL10M3-F50-EEE, DSI) which allows the simultaneous recording of cortical and hippocampal (deep into the *dentate gyrus*) electroencephalograms and the electromyogram EMG (on the neck muscle) [6]. After approximately 45 days of post-surgery recovery, rats were singly housed and habituated for 5 days into the experimental PhenoTyper cages (by Noldus Information Technology). After this relatively short habituation period, a background registration session was conducted acquiring simultaneously and continuously for 24 hours the telemetric signals using the acquisition system Data Science International Dataquest ART Gold 4.01, and the general activity of animals by collection of videos using the MPEG4 Encoder and The Observer XT connected directly with the PhenoTyper cages. Three days after this background recording, two further sessions were conducted, in which 2 out of the 4 animals received a novel environmental enrichment (3 objects: tunnel, ball and rodent house) into their cages. The first recording's session was conducted immediately and the second three days after the novel environmental enrichment. Three days later, two additional sessions of recording were conducted, in which for 4 consecutive days at approximately 8.00 am all 4 animals received an olfactory stimulus (3 drops of bergamot and 3 of lemon essential oil applied into the bedding material – common perfumes exhaled by disinfectants used for cleaning procedures) into their cages; the first recording's session was conducted immediately after the first occasion of the olfactory stimulus and the second on the fourth day of olfactory stimulus.

On each recording's occasion the telemetric signals and the videos were acquired simultaneously and continuously for 24 hours, starting at approximately 8.00 am and stopping at the same time (8.00 am) of the day after.

Telemetric EEGs, EMG waveforms and the activities of animals collected with our integrated telemetric and video system will be presented thus providing unprecedented insight into the correlations between behavioural and brain electrical activities following environmental and olfactory experiences.

All the experiments were carried out in accordance with Italian regulation governing animal welfare and protection and the European Directive 86/609/EEC, and according to internal GlaxoSmithKline Committee on Animal Research & Ethics (CARE) review.

References

1. Van Praag, H., Kempermann, G., F. H. (2000). Neural consequences of environmental enrichment. *Nat. Rev. Neurosci.*, **1**, 191-198.
2. Mora F., Segovia G., Del Arco A. (2007). Aging, plasticity and environmental enrichment: Structural changes and neurotransmitter dynamics in several areas of the brain. *Brain Res. Rev.*, **55**, 78-88.
3. Lucero M. (1970). Leghtening of REM sleep duration consecutive to learning in the rat. *Brain Res.*, **20**, 319-322.
4. Mirmiran M., Van den dungen H., Uylings H.B M. (1982). Sleep patterns during rearing under different environmental conditions in juvenile rats. *Brain Res.*, **233**, 287-298.
5. Zibrowski E.M., Vanderwolf C.H. (1997). Oscillatory fast wave activity in the rat pyriform cortex: relations to olfaction and behaviour. *Nat. Rev. Neurosci.*, **766,** 39-49.
6. Williams P.A., White A.M., Ferraro D.J., Clark S., Taley K. and Dudek F.E. (2006). The use of radiotelemetry to evaluate electrographic seizures in rats with kainite-induced epilepsy. *J. Neurosci. Methods*, **155**, 39-48.

Automated phenotyping using intra-home-cage technology

Y.K. Urbach[1], K.A. Raber[1], T. Appl[1], L. Häberle[2], and S. von Horsten[1]
[1]*Friedrich-Alexander-University of Erlangen-Nürnberg, Franz-Penzoldt-Center, Experimental Therapy, Germany, yvonne.k.urbach@ze.uni-erlangen.de*
[2]*Institute for Medial Informatics, Biometrics, and Epidemiology (IMBE), Friedrich-Alexander-University Erlangen-Nürnberg, Germany*

Similar to other approaches into this direction, the PhenoMaster system (TSE Systems, Germany) is capable to automatically screen rats in a home-cage-like environment for several parameters at a high temporal and spatial resolution. This novel approach generally allows experimenter independent monitoring of laboratory rodents and – a priori – allows avoidance of stress-artifacts, higher throughput, higher sensitivity of measures (online, circadian), and combination of measurements potentially allowing multifactorial analysis and idenification of novel combined behavioral dimensions. Especially, the newly developed PhenoMaster System for rats is conceived as an automated modular high throughput screening system for the assessment of specific gene-associated functions on the physiological and behavioral phenotype of small laboratory animals. It is a modular, multi-purpose tool for investigation of multidimensional physiobehavioral outputs of rats.

Apparatus PhenoMaster

The PhenoMaster system represents a modular set-up that measures indirect calorimetric parameters, activitiy, drinking, feeding, operant wall or wheel and is based on conventional type IV Thermoplast cages, with each cage being equipped with the technical requirements to individually monitor one rat per cage at a time. The present system is set-up to measure 12 animals in parallel with high resolution for activity, calorimetric parameters and food and water consumption. The animals are monitored for the following primary measures: food and water intake, RER, and locomotor activity. Secondary measures are: energy expenditure and other combined readouts. Cumulative feeding and drinking were given and this was used to compare iterative food- and water intake. The calorimetry system is an open-circuit system that determines O_2-consumption, CO_2-production, and RER. To investigate locomotor activity, a photobeam-based activity monitoring system detects and records the number and duration of every total, fine and ambulatory movement, including rearing and climbing movements, in every cage.

This activity detection is achieved using infrared sensor pairs arranged in strips for horizontal (x, y level) and vertical (z level, rearing) activity. Light beams with 32 beams in X, 25 beams in Y dimension and 32 beams in the Z-axis are used. The mean activity per hour (in 20 min intervals) over 72h for each parameter is calculated. Results for locomotor activity are given as number of counts. In addition, depending on the pattern of interruption of the light beams, movements are further subdivided by the PhenoMaster software into fine movements (XF, YF; "grooming") defined as the counting of repeated beam breaks of the same light barrier, ambulatory movement that are the counting breaks of alternate barriers (breaks of Y-axis, XA, YA) and peripheral ambulatory movements (PerT, PerA, PerF). In addition, the total number of interruption of the Z-axis was monitored ("rearing"). Only interruptions of the light beams that are linked with a movement will be recorded, therefore it is possible to define a refractory period so that short-term movements can be ignored and permanently interrupted light beams will not be taken into account as well.

The sensors for detection of movement operate efficiently in both light and dark phases, allowing continuous recording. All measurements are monitored using the PhenoMaster software.

Presently, the system is used to characterize changes in energy-related parameters in rats transgenic for Huntington's disease and Spinocerebellar Atxia type 17 in the course of the European-project RATstream™ (www.ratstream.eu).

Experiments were done in a separate room under standard conditions with control units and computer within the same room. Subsequently, animals were introduced to the system and parameters were continuously and simultaneously measured during 72h. During this period of testing the experimenter only once a day entered the room for controlling proper working of the system and checking health of animals by visual inspection.

Analysis of data

Energy expenditure was measured as VO_2 consumption in ml/(h*kg/BW) and VCO_2 production in ml /(h*kg/BW) and the respiratory quotient (RER) was calculated as the quotient of VCO_2 production divided by VO_2 consumption for each 20 min. Cumulative feeding and drinking was taken as a sum of consumed food- and water intake per 20 minutes iterative. Locomotor activity counts were summed for 20 min intervals from measurements over 1 min intervals for the duration of the experiment. Mean activity counts were calculated for each experimental group and analyzed as activity over time. Raw data derived from automated PhenoMaster system was exported for statistical analysis using R system for statistical computing (version 2.4.1; R Development Core Team, Vienna, Austria; 2006). Imports were examined graphically (Q-Q plots, histograms, boxplots) to determine if normal distribution and equal variances can be assumed for each parameter. For normal distribution, one-way ANOVA with one between-subject factor (genotype) and with repeated measures over time was performed. If data are not normally distributed, non-parametric tests for repeated measures according to the method of O'Brien were applied.

Assessment of individual differences in behavioural inhibition in mice across different testing paradigms and in a homecage-test situation (IntelliCage)

F. Magara[1], Th. Steimer[2], and J. Ackermann[3]

[1]Center for Psychiatric Neuroscience, Dept of Psychiatry, Lausanne University Hospital, Switzerland, Fulvio.Magara@unil.ch

[2]Clinical Psychopharmacology Unit, Geneva University Hospital, 1225 Chêne-Bourg, Switzerland, Thierry.Steimer@hcuge.ch

[3]Dept of Cell Biology and Morphology, Lausanne University, Lausanne, Switzerland, Julien.Ackermann@unil.ch

Mice of the strain 129X1/SvJ, a major ES cell donor, are considered to be dull due to their passivity and poor performance on a variety of behavioral paradigms. It has been proposed, however, that this hypoactivity is due to a neophobic reaction to the test settings, rather than to a limited cognitive ability(Homanics *et al.*, 1999), (Dockstader & Van der Kooy, 2001)

We run a preliminary series of tests (elevated plus maze, Làt maze, light/dark box) on 55 mice, we found a bimodal distribution of exploratory measures within a cutoff of 10 min, with 30% to 60% of the mice strongly inhibiting locomotion in any arena. When re-exposed to the same testing arena, most mice consistently replicated their locomotor responses, albeit a sort of inverse habituation (i.e. increase of activity upon retesting) was observed for some of them. Measures of locomotor activity appeared to be consistent both within and across arenas, mice refraining to move in a given setting being also hypoactive in other settings.

In the attempt of generating two recombinant lines differing for their coping attitudes we bred to each other mice selected for high (respectively low) latencies to enter the novel compartments in a free exploration paradigm (FEP) (Misslin *et al.*, 1982). The FEP consists of a grey PVC box divided in 6 compartments, measuring 10x10 cm. The mouse is allowed to habituate to three communicating compartments during 24 hours; thereafter, access to three adjacent compartments is made available by opening of guillotine doors. Latency to enter the novel compartments, percent of time spent in the novel side and risk assessment events are measured on a 10 minute time lapse from door opening.

This psychogenetic selection produced so far two F3 generations quite differing in locomotor activity, rather than on measures of anxiety, as assessed in the open-field, elevated plus maze and FEP. Quantification of locomotor tracks produced by EthoVision revealed similar between-session habituation in the two groups of mice, yet different within-session habituation profiles, suggesting that the psychogenetic selection specifically affected locomotion, or a form of anxiety, yet not memory abilities. Indeed, measures of anxiety taken from different tests did not correlate to each other; however, activity did, the less active mice moving poorly on virtually any arena.

Exploratory attitudes appear to correlate inversely to the stress response: ten minutes after door openings in the FEP, mice were either (1) allowed to continue exploration of the novel compartments, or (2) re-confined in the familiar compartments, or (3) forced and confined in the novel compartments. Afetr 10 minutes, blood was samples from the tail vein and plasma corticosterone assessed by RIA. Results show that corticosterone levels were inversely correlated to the amount of locomotion in response to the opening of the doors.

To understand whether this locomotor inhibition had to be ascribed to a form of neophobia, we studied the exploration and place learning patterns of F3 inhibited and non-inhibited mice in the Intellicage™.

The Intellicage is a sort of automated operant conditioning meant to evaluate associative learning in a social, homecage setting: the access to water bottles, set in the corners of the cage, is gated by a tunnel and doors, and can be programmed in such a way to impose an operant schedule of visits and nosepokes to each individual mouse, recognized by means of transponders. Because "wrong" responses can also be "punished" with an air puff, the device promises application also in the study of avoidance learning and fear response. We put 4 inhibited and 7 non-inhibited mice together in two Intellicages, and programmed the corners in such a way that half of the mice (2 inhibited and 3 non-inhibited) received an airpuff when visiting corner 1, while the other half (2 inhibited and 4 non-inhibited) received an airpuff when visiting corner 3. Corners 2 and 4 gave free access to water to all mice.

Latencies to first visit the corners of Intellicage are higher for the F3 offspring of inhibited mice: the same animals, however, readily learn to avoid punished corners. By 3 days, all mice habituate and differences in activity between inhibited and non-inhibited animals become undetectable. Surprisingly, all mice, no matter whether punished in corner 1 or 3, learned to avoid both punished corners. Apparently, inhibited mice learned even faster than the non-inhibited to avoid the punished corners.

These results suggest that the poor proficiency of the 129X1 mice is largely due to a form of trait anxiety that results in a strong locomotor inhibition, rather than to impaired learning and memory abilities.

In order to further assess the effect of the testing environment vs homecage on measures of anxiety, the study has been extended to a 5HTR1 mutant mouse line, where elements of locomotor inhibition make difficult the interpretation of measures of anxiety.

References

1. Dockstader, C. L. & Van der Kooy, D. (2001). Mouse strain differences in opiate reward learning are explained by differences in anxiety, not reward or learning. *Journal of Neuroscience*, **21**, 9077-9081.
2. Homanics, G. E., Quinlan, J. J., & Firestone, L. L. (1999). Pharmacologic and behavioral responses of inbred C57BL/6J and strain 129/SvJ mouse lines. *Pharmacol.Biochem.Behav*, **63**, 21-26.
3. Misslin, R., Herzog, F., Koch, B., & Ropartz, P. (1982). Effects of isolation, handling and novelty on the pituitary--adrenal response in the mouse. *Psychoneuroendocrinology*, **7**, 217-221.

Dissecting and validating the salience of recognition cues used by female zebra finches' to discriminate con- and heterospecific males

D.L.M. Campbell and M.E. Hauber

School of Biological Sciences, University of Auckland, Auckland, New Zealand, d.michelle@auckland.ac.nz

Successful sexual reproduction relies on accurate mate selection criteria, including correct species recognition [1, 2]. In studying species recognition in birds, video playback of visual and/or acoustic cues can be a valuable method for investigating the relative importance of these sensory cues in the recognition process by allowing independent manipulation of the two communication modes. Despite video-playbacks' widespread use to disassociate the roles of visual and acoustic recognition cues, the validity of behavioural responses towards video-playbacks have not yet been tested in a free-choice pair bonding paradigm. In a model system for social recognition research, the zebra finch (*Taeniopygia guttata*) [3], we investigated females' behavioural responses to playbacks of male conspecific and heterospecific cues from an experimentally novel sympatric species, the plumhead finch (*Neochmia modesta*). Multiple unfamiliar male conspecifics and heterospecifics were individually filmed and footage was edited to create video playback stimuli of a male singing and performing additional courtship behaviours. Simultaneous (but acoustically alternating) playbacks were presented to the female subjects. To investigate the relative importance of visual and acoustic cues in species recognition we presented playbacks of visual cues only, acoustic cues only, these cues concurrently and then a combination of visual and acoustic cues where the songs of the two stimulus species were swapped and edited so that both species appeared to be visually singing each others song. This study represents the first application of the use of heterospecific song in species recognition in female zebra finches and the first to alter the visual and acoustic cues without affecting the spatial association of these cues that are used for proximity decisions [4]. The female subjects' response behaviour was recorded using a scan-sampling design every 10 s where both location and behaviour was noted. The species discrimination preferences of females were determined by calculating the average proportion of total choice time that they spent in proximity to the conspecific (or conspecific acoustic when the songs were swapped between species) [3]. In addition to spatial discrimination we measured and found significant individual variation in behavioural response to the video playbacks themselves irrespective of presented species.

In monogamous species such as the zebra finch where both sexes invest in parental care there is minimal reciprocal research on criteria that males may use for selecting females [5, 6, 7]. To validate our behavioural assay, we predicted that significant behavioural variation displayed by female zebra finches in our employed testing context might be able to correlate with mating decisions made by male zebra finches. By observing pair bonding behaviours in a free-flight aviary mate choice paradigm, the relevance of females' variation in mate selection behaviour (choosiness) [8] in our experiments to behavioural pairbonding was confirmed. We found significant positive correlations between the average proportion of total choice time that females were actively engaged with the visual presentation of males and male zebra finches' preferences for a particular female (see Figure 1). These results are the first to demonstrate a biologically relevant statistical relationship between variable female choice behaviours in one testing context and subsequent mate choice decisions made by males using an alternative testing paradigm in a socially monogamous species.

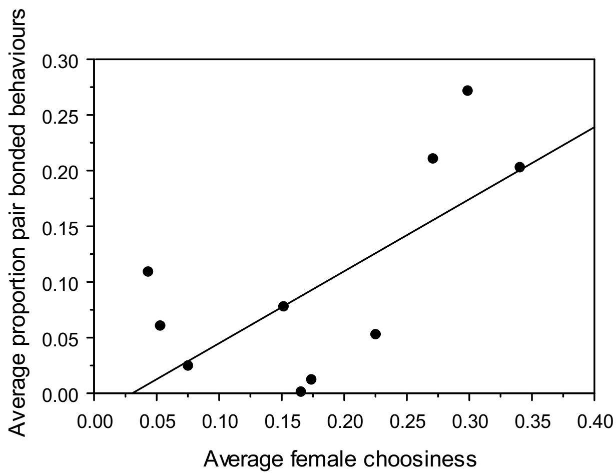


Figure 1. *A regression plot showing the average proportion of total pair bonded behaviours that each female was selected to engage in with the subject males in relation to the average observed female choosiness in the video playback trials. Individual data points indicate the 10 female stimuli.*

References

1. Irwin, D.E., Price, T. (1999). Sexual imprinting, learning and speciation. *Heredity*, **82**, 347-354.
2. Price, T. (2008). *Speciation in Birds*. Colorado: Roberts and Company Publishers.
3. Zann, R.A. (1996). The Zebra Finch: A synthesis of field and laboratory studies. USA: Oxford University Press, Inc.
4. Brazas, M.L., Shimizu, T. (2002). Significance of visual cues in choice behaviour in the female zebra finch (*Taeniopygia guttata castanotis*). *Animal Cognition*, **5**, 91-95.
5. Jones, K.M., Monaghan, P., Nager, R.G. (2001). Male mate choice and female fecundity in zebra finches. *Animal Behaviour*, **62**, 1021-1026.
6. Monaghan, P., Metcalfe, N.B., Houston, D.C. (1996). Male finches selectively pair with fecund females. *Proceedings of the Royal Society of London*, **263**, 1183-1186.
7. Ikebuchi, M., Okanoya, K. (2006). Growth of pair bonding in zebra finches: physical and social factors. *Ornithological Science*, **5**, 65-75.
8. Jennions, M.D., Petrie, M. (1997). Variation in mate choice and mating preferences: a review of causes and consequences. *Biological Reviews*, **72**, 283-327.

The ultraviolet (UV) irradiation Induced Heat Hyperalgesia rat biomarker model for Pain Research

Thomas Pitcher, Ian Machin, and Laura Corradini
Pfizer R&D, Pain Therapeutics, Sandwich, Kent, UK, tom.pitcher@pfizer.com

The exposure of the skin to Ultraviolet (UV) irradiation results in a classical inflammatory reaction characterised by both erythema (flare) and hypersensitivity to noxious and non noxious stimuli (i.e. hyperalgesia and allodynia). No spontaneous pain is associated to this injury [1,2].

Cutaneous UV irradiation causes DNA damage, activates MAP kinase in keratinocytes, which leads to increased levels of pro-inflammatory cytokines, such as TNFα and IL-8, and up regulation of COX-2 [3]. Pain sensitivity is increased after thermal stimulus in the area of erythema (primary hypersensitivity).

This inflammatory reaction has been studied in both humans and rodents and represents a potential translatable pharmacology biomarker model for the assessment of efficacy of new analgesics [1, 4-6].

Methods

Animals- All procedures were carried out under the Animals (Scientific Procedures) Act, 1986. Male Sprague Dawley rats (200-300g, Charles River, UK) were housed in groups of 4 under a 12 hour light/dark cycle with food and water *ad libitum*. All experiments were carried out by an investigator blind to drug treatments.

UV Radiation Exposure

Animals were anaesthetized with a 2% isoflorane 0_2 mixture. Anesthesia was maintained via a nose cone while the plantar surface of the right paw was irradiated with 100, 200, 300 or 400mJ/cm^2 of UV irradiation. This exposure was generated by Saalmann CupCube system (Saalmann GmbH, Germany; λ=280-400 nm). The source of irradiation was adapted to encompass the plantar surface of the rats hind paw by using a shaped delivery collar (either 8mm diameter or 8x12mm oval), attached to the UV emitter (Pfizer Facilities Management and Engineering Team). The intensity of UV emitted was calibrated prior to exposure using an ABLE 1400A radiometer with SEL005/WBS320/TD filter (ABLE Instruments and Controls Ltd).

Thermal hyperalgesia

Heat hypersensitivity was assessed using the rat plantar test equipment (Ugo Basile, Italy) following a modified method as described by Hargreaves et al (1988). The plantar test consisted of three Perspex boxes (22x19x25cm) on an elevated glass table. Two rats were housed in each box, so that 6 rats could be tested simultaneously in a single apparatus, and left to acclimatize for at least 20 mins. A mobile infrared heat source was applied to the plantar surface of the hind paws. The paw withdrawal latency (PWL) was defined as the time (in seconds) taken by the rat to remove its hind paw from the heat source. The heat source was calibrated to give a response of 8-12 sec on untreated animals (105-110mW/cm^2). An automatic cut off point of 20 sec was applied to prevent tissue damage. Uninjured paw was always assessed first. Five recordings (1-2 mins intervals between recordings) from each paw, were taken during baseline recordings and expressed as a normalised mean (highest and lowest value removed). In the pharmacology test session, three recordings were collected from each paw, the average of which represents PWL for each rat.

Statistical analysis

Data are expressed as mean of 6-8 rats per group. Each treatment group was compared at each time point to vehicle-treated group using a one way ANOVA followed by a Dunnets t test, blocked for each day of treatment.

Results

The time course of development of thermal hyperalgesia was evaluated as well as the activity of a number of standard agents seen effective in chronic pain diseases.

A cohort of Sprague Dawley rats was exposed to various intensities of UV irradiation (100-400 mJ/cm^2; n=6-9 rats/group) and thermal hyperalgesia was assessed every day up to 7 days post UV application.

The development of hypersensitivity showed a significant reduction in thermal PWL from 24hrs post UV in all groups. The peak of hyperalgesia was observed at 48 hours, maintained up to 96 hours but fully resolved at 7day post the induction of erythema. Within the range of UV exposures used, 300 mJ/cm2 appear to produce the highest degree of hypersensitivity without causing tissue damages. Dry skin was indeed observed with higher intensity. No vesicles or severe skin damage was seen at any of the exposures used.

The efficacy of various standard analgesic therapies was assessed. The non selective, COX inibitor, Ibuprofen (100 and 300 mg/kg PO), the COX2 inhibitor, Valdecoxib (30mg/kg, PO), the μ-opioid agonist, oxycodone (0.3-1 mg/kg, SC) and the non selective sodium channel blocker, Mexiletine(10-30 mg/kg, SC) significantly reversed hyperalgesia at various time points post drug administration and data were successfully reproduced by two investigators in double blind study design.

Conclusion

We have demonstrated that post UV skin exposure, rats develop thermal hyperalgesia localized at the area of erythema. This hypersensitivity can be pharmacologically reversed by standard compounds seen therapeutic in inflammatory (Ibuprofen), chronic nociceptive (Oxycodone and Valdecoxib) and neuropathic pain condition (Oxycodone and Mexiletine). Most of this mechanisms have been able to reverse heat hypersensitivity in the human UV model [4,7; internal data) so that it has been proposed this model could be a useful translatable pharmacology biomarker for Pain research.

Further analysis is on going to explore whether UV model could also be a suitable model to escalate efficacious doses from rat to man.

References

1. Bishop T, Hewson DW, Yip PK, Fahey MS, Dawbarn D, Young AR and McMahon SB. (2007). Characterisation of ultraviolet-B-induced inflammation as a model of hyperalgesia in the rat. *Pain*, **131**, 70-82.

2. Davies SL , Siau C and Bennett GJ (2005). Characterization of a model of cutaneous inflammatory pain produced by an ultraviolet irradiation-evoked sterile injury in the rat. *Journal of Neuroscience Methods*, **148**, 161-166.
3. Hruza LL and Pentland AP (1993). Mechanisms of UV-induced inflammation. *J. Invest. Dermatol.*, **100**, 35S-41S.
4. Gustorff B, Hoechtl K, Sycha T, Felouzis E, Lehr S, Kress HG. (2004). The effects of remifentanil and gabapentin on hyperalgesia in a new extended inflammatory skin pain model in healthy volunteers. *Anesth Analg*., **98**, 401-7.
5. Harrison GI, Young AR and McMahon SB. (2004). Ultraviolet Radiation-Induced Inflammation as a Model for Cutaneous Hyperalgesia . *J. Invest. Dermatol*., **122**, 183-189.
6. Hamilton SG, Wade A and McMahon SB (1999). The effects of inflammation and inflammatory mediators on nociceptive behaviour induced by ATP analogues in the rat. *Br. J. Pharmacol.*, **126**, 326-332.
7. Sycha T , Anzenhofer S, Lehr S, Schmetterer L, Chizh B, Eichler HG and Gustorff B (2005). Rofecoxib attenuates both primary and secondary inflammatory hyperalgesia: a randomized, double blinded, placebo controlled crossover trial in the UV-B pain model. *Pain*, **113**, 316-322.

The behavioural differences between C57B6L/6 and 129Sv mice are reproducible in mice reared in distinct environmental conditions

Urho Abramov, Triinu Puussaar, Sirli Raud, Kaido Kurrikoff, and Eero Vasar
Department of Physiology, University of Tartu, Estonia, Urho.Abramov@kliinikum.ee

Introduction

In a study initially published by Crabbe *et al* (1999) and later discussed in more detail by Wahlsten *et al* (2003) a major conclusion was drawn, that the behavioural differences with large effect size between the mouse strains could be reproduced in different laboratories, whereas the differences with small effect size would be more prone to modification by laboratory environment. We have recently reported that the phenotype of mice, lacking cholecystokinin CCK-2 receptors, could not be reproduced within our laboratory if mice had been reared in different housing conditions (Abramov et al, *in press*), partially explaining large number of contradictory reports on these mutants over past 10 years. Provided that the mice, lacking CCK-2 receptors, did not significantly differ from their wild-type littermates in terms of genetic background, the present study was designed to reveal if the behavioural phenotypes of genetically distinct strains could be reproduced in mice reared in distinct environmental conditions within our laboratory. Environmental enrichment was used as an alternative environment to standard laboratory conditions as suggested by Würbel (2002).

Methods

Studies were performed in male and female C57Bl/6 (B6, Scanbur BK) and 129S6/SvEv/Tac (129, Taconic) mice. After weaning at 3 weeks mice were randomly allocated to either standard or enriched conditions for 7 weeks before start of experiments. Standard housing conditions consisted of standard laboratory cages with bedding, whereas enriched conditions consisted of larger cages containing bedding, nesting material, stainless steel wheels or swings and aspen houses, igloos, ladders, tubes or labyrinths, which were changed and repositioned once a week. For behavioural phenotyping the plus-maze, locomotor activity, hot-plate and forced swim test were carried out (for details, Abramov *et al*, *in press*).

Results

Significant genotype-dependent differences were established in all tests. 129 strain displayed significantly lower exploratory activity in the plus-maze (Fig. 1) and locomotor activity test (Fig. 2) than B6 strain. Also, significantly longer latencies to hind-paw reaction were observed in 129 strain in the hot-plate test when compared to B6 strain (Fig. 3), and 129 strain spent substantially larger proportion of time immobile than their B6 counterparts in the forced swim test (Fig. 4). With the exception of the plus-maze test, the differences between B6 and 129 strains were significant, irrespective of pre-experimental housing conditions. Interestingly enough, behavioural differences between strains were even more pronounced in mice reared in enriched conditions.

Conclusions

The present study demonstrates, that 129 strain displays lower exploratory activity, longer hot-plate latencies and spends more time immobile in the forced swim test when compared to B6 strain. These behavioural differences were reproducible in mice reared in distinct environmental conditions, indicating clear genotype-dependent effects rather than genotype by environment interaction effects. In the light of previous findings (Abramov et al, *in* press), these results suggest, that rearing in two distinct environmental conditions is a meaningful approach in behavioural research and can be applied to dissect a phenotype into effects arising from genes, environment or the combination of both.

References

1. Abramov U et al. (*In press*) Different housing conditions alter the behavioural phenotype of CCK_2 receptor deficient mice. *Behav Brain Res*.
2. Crabbe JC et al. (1999) Genetics of mouse behavior: interactions with laboratory environment. *Science*, **284**, 1670-1672.
3. Wahlsten D et al. (2003) Different data from different labs: lessons from studies of gene-environment interaction. *J Neurobiol*, **54**, 283-311.
4. Würbel H. (2002) Behavioral phenotyping enhanced – beyond (environmental) standardisation. *Genes Brain Behav*, **1**, 3-8.

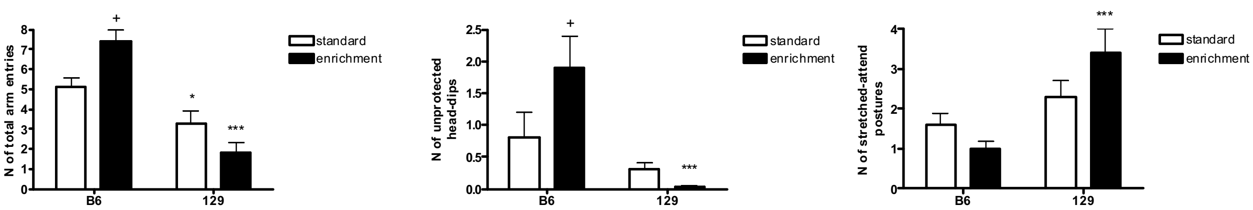


Figure 1. *Plus-maze test in B6 and 129 strains, housed in different environmental conditions. (A) total number of arm entries; (B) number of unprotected head-dips; (C) number of stretched-attend postures. * p<0.05, *** p<0.005: 129 strain compared to B6 strain housed in the same conditions; + p<0.05: mice housed in enriched conditions compared to mice of respective strain housed in standard conditions.*

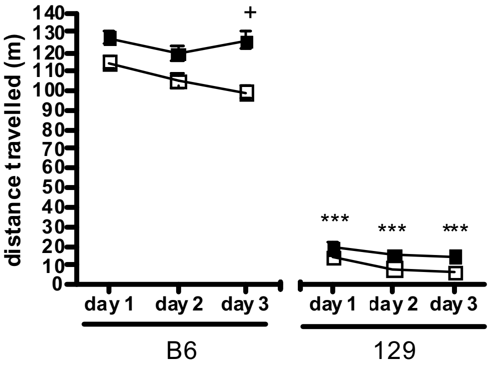


Figure 2. *Locomotor activity in B6 and 129 strains, housed in different environmental conditions. *** p<0.005: significant differences between 129 and B6 strains on the same day irrespective of housing conditions; + p<0.05: B6 strain housed in enriched conditions compared to B6 strain housed in standard conditions on day 3.*

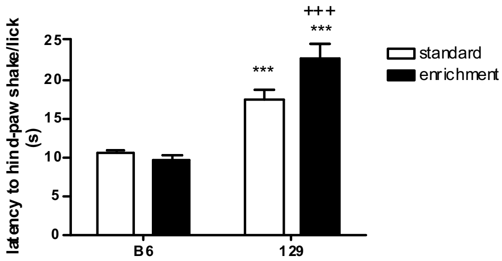


Figure 3. *Hot-plate test in B6 and 129 strains, housed in different environmental conditions. *** p<0.005: 129 strain compared to B6 strain housed in the same conditions; +++ p<0.005: 129 strain housed in enriched conditions compared to 129 strain housed in standard conditions.*

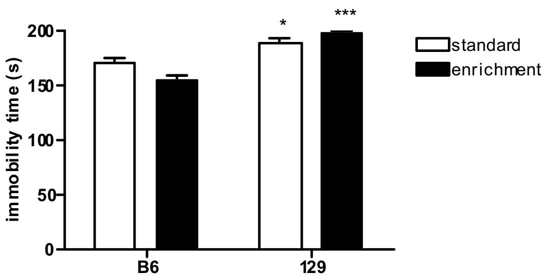


Figure 4. *Forced swim test in B6 and 129 strains, housed in different environmental conditions. * p<0.05, *** p<0.005: 129 strain compared to B6 strain housed in the same conditions.*

Inter- and intra-observer reliability with data recorded through '*The Observer*' and '*Match Vision Studio*'

Ángel Blanco-Villaseñor and M. Teresa Anguera
Department of Behavioural Sciences Methodology, University of Barcelona, Barcelona, Spain
ablanco@ub.edu, mtanguera@gmail.com

Controlling data quality is highly important in any study based on observational methodology and it comprises three key aspects: reliability, accuracy and validity (which may be addressed together in a generalizability study). There are at least three ways of addressing the reliability of observational data: 1) coefficient of concordance between the observer's judgments (the agreement between them), which refers to the observations made by the different observers at a given point in time; 2) an observational measure may be considered as a special case of a standardised psychological test, and thus we can use the definitions of reliability associated with classical psychometric theory, i.e. the correlation coefficient; and 3) an observational measure may yield data under the influence of a given number of different aspects of an observational situation (different observers, different occasions, different ways of recording, different recording instruments), including the individual differences between participants. This third aspect concerns the GT developed by Cronbach, Gleser, Nanda & Rajaratnam (1972).

One good way of conducting the whole recording process, data quality control and data analysis is to use the data output from *The Observer* and *Match Vision Studio* software, since the output is generated via an Excel spreadsheet. A further advantage is that the data can then be readily transformed from one format to another. Thus, this source of data may be used directly for all the calculations, regardless of whether they concern reliability or data analysis. For example, they may be transformed into the CSV format and analysed using Theme.

A computer based rating system for testing motor performance in children

H. Holzer[1], O. Fleiss[2], H. Fleiss[2], and S. Kellner[1]
[1]*Institute of Sports Science, Karl-Franzens-University, Graz, Austria, hans-peter.holzer@uni-graz.at*
[2]*Steirische Gesellschaft für Wirbelsäulenforschung, Graz, Austria*

Introduction

Motor fitness tests of children show that up to 35% of the children have a poor motor fitness at schoolstart [5]. In our own studies about posture problems of schoolchildren we found that up to 40% of the children also had coordination problems [1,2]. The basis of our findings were videos which we had taken of posture, spinal movement and some overall body movements, that allow an insight into the diversity of movement behavior. In order to evaluate the videos in regards to coordination problems, we use a computer based rating and documenting system that allows to analyze the main individual movement characteristics.

Methods

Our rating system is based on the performance of six basic movements: walking, running, standing and jumping on one leg, side jumping on both legs and a jumping jack (see Figure 1). These movements are filmed by video.

In order to carry out the ratings the film sequences (test-movements) along with (up to 10) pictures (special positions for documentation of movement problems) are transferred to the computer.

For each test-movement, more than 80 characteristics can be rated. These characteristics are grouped into five main-parameters: *personality in motion, posture, movement-pattern, quality of movement, coordination, coordinative abilities, and motor-characteristics*. An additional group documents *special signs* such as a fixated or rotating arm or extra movements (see Figure 2).

The rating takes place on a four point scale but in contrast to most of the test systems of today the rating takes place by analyzing movement parameters - *and* posture parameters. In doing this each chosen parameter contributes to the rating of a single characteristic, the associated main parameter and the coordination as a whole. This means you don't have to judge a movement by counting how often one can perform it in a given time or by assessing the quality of the movement. For example by choosing the "inward rotated foot" (test-movement *walking*) of the associated menu (foot) a weighted contribution of "minor problems" to the main parameter *posture* is made. A weighted sum of all the rated parameters then allows the quantification of the motor performance itself.

In order to achieve an exact individual analysis, all items that apply to the specific movement should be considered. This is a time consuming process, so there is also a "short-cut" incorporated into the program. In most cases a rating of up to 10 relevant parameters per test-movement is sufficient. These parameters we call *checkpoints* (see Figure 2). By choosing a test movement these *checkpoints* are displayed and allow an efficient way to analyze a motor performance.

The coordination profile documents the intra-individual differences regarding the different test movements and the evaluation of the main parameters as mentioned above. Additional an individual "strength-weakness" profile of the tested child is available.

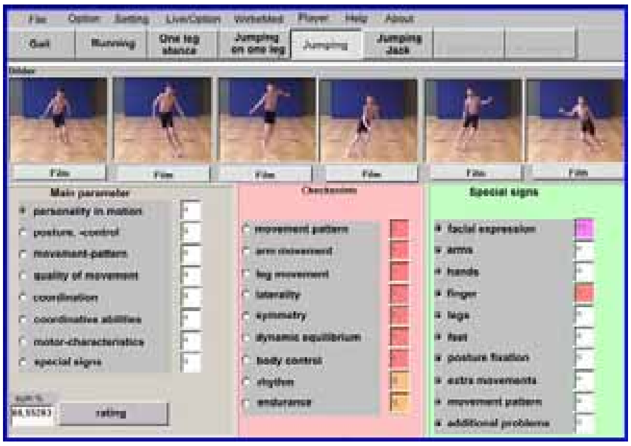


Figure2. *Rating window - jumping on both legs.*

Results & Discussion:

Using this system, a study was done to evaluate the motor performance of Styrian schoolchildren (N=20) from the ages of six to ten years. Coordination weaknesses were found in 55% of the children whereas only 25% showed excellent coordination. Coordination problems were found in only 5%. As expected, such weaknesses were frequently found in the complex exercises jumping jack (see Figure 3) and jumping on both legs.

Figure 1. *Test-movements.*

***Figure 3**. Jumping Jack – coordination problems*

In this first field-test, this system proved to be an effective tool for the examination and description of motor coordination. The findings (as far as comparable) were in good correlation with data from literature [3-6]. The interaction of the various parameters will be part of future tests. An additional help function by means of a film and picture database with case studies is being prepared, as well as special exercise programs for children with coordination problems.

Conclusion

With this coordination testing system, it is possible to evaluate individual characteristics as well as the testing of the overall motor performance. Thus, one can take care either of individual weaknesses or support good motor abilities.

References

1. Holzer, H., Fleiss, O., Fleiss, H., Ritter G. (2000). Computerunterstützte Haltungsuntersuchung bei Schulkindern. Prävention und Rehabilitation des Haltungs- und Bewegungsapparates: *DVS Schriftenreihe*, **103**, 179-185.
2. Holzer, H., Fleiss, O., Fleiss, H. (2002). Computer-aided examination of schoolchildren´s postural faults. Proceedings of Measuring Behavior 2002, 4th International Conference on Methods and techniques in Behavioral Research (Amsterdam), 105-106.
3. Ketelhut, K. (2000). Bewegungsmangel im Kindesalter: Gesundheit und Fitness heutiger Kinder besorgniserregend?. *Dt. Zs. Sport Med.*, **51(10)**, 350.
4. Sönnichsen, C., Weineck, J., Köstermeyer, G. (1997). PEP eine Studie zur Präventionserziehung. Teil 3: Motorisches Leistungsvermögen und kardiovasculäres Risiko von Schulanfängern. *Haltung Bewegung*, **3**, 5–10.
5. Starker, A., Lampert, T., Worth,A., Oberger, J., Kahl, H., Bös, K. (2007). Motorische Leistungsfähigkeit:Ergebnisse des Kinder und Jugendsurveys (KIGGS). *Bundesgesundheitsbl - Gesundheitsforsch -Gesundheitsschutz*, **50**, 775–783.
6. Weineck, J., Köstermeyer, G., Sönnichsen, C. (1997). PEP eine Studie zur Präventionserziehung – Teil 1: Zum motorischen Leistungsvermögen von Schulanfängern. *Haltung Bewegung*, **17**, 5–16.

Tracing motivation in multimedia learning contexts

M. Zhou and P.H. Winne
Faculty of Education, Simon Fraser University, Burnaby, Canada, mzhou2@sfu.ca

More recently, theorists have focused on achievement goals, conceptualized as situationally specific measures of motivational orientation, and argued that they could be stronger predictors of academic success [1]. Current models of self-regulated learning integrate motivational and cognitive elements of learning, showing how achievement goals and learning expectancies influence students' use of cognitive processes [2-4]. The basic assumption of these models is that the achievement motives and the intentions that guide students' academic behavior determine to a great extent the types of cognitive processes they employ in various learning situations. The learning outcome is thus dependent on how students process information [5, 6].

The relationships between goals, strategic behavior, and performance are not, however, straightforward and consistently supported by research. One of the focal issues is a consideration of how the learning context influences student motivation and self-regulation. Multiple earlier empirical findings have shown that students may adopt context-specific interpretations of motivational goals and self-regulation [7, 8]. More recent work indicates incongruence between general level goal orientations and contextual goal interpretations and, thus, underlines the importance of subjective interpretations of achievement goals (e.g., [9, 10]). This contextual characteristic illustrates the need to understand the interplay of motivational and cognitive processes.

Traditional Methods in Research on Motivation

Most of the research evidence on students' motivational goal orientations and learning strategies has been obtained from studies using self-report questionnaires or interviews as the primary source of data. The most fundamental issue with most self-report measures seems to be that they do not take the fact that students modify their strategic processing to fit different tasks and purposes into account (e.g., [11]). Further, they do not permit an in-depth exploration of how individuals come to construct their understanding within a learning environment or the role of the contextual factors [12].

Retrospective interviews have been suggested as useful methods since they allow respondents to reveal and explain events and experiences in detail [13], and could be structured in a way that respondents describe their behavior in a particular situation in which the interview is anchored [14]. Nonetheless, answering questions is not a simple process but consists of various stages of complex cognitive processing [15] in that the respondent has to interpret and comprehend the question, retrieve relevant information from autobiographical memory, use heuristic and other decision-making processes to estimate an answer, and then formulate a response [16]. Meanwhile, memory is subject to loss (forgetting), distortion (biased sampling of memories), and reconstruction during this cognitive processing. All these effects can undermine the accuracy and reliability of responses.

Tracing Method with Log Data

Recently, the importance of reaching students' situational interpretations in real on-line learning situations has been highlighted [17]. gStudy (Figure 1) is a software application for researching self-regulated learning [18] in solo and collaborative settings. Learners use gStudy to operate on multimedia content with a variety of cognitive tools.

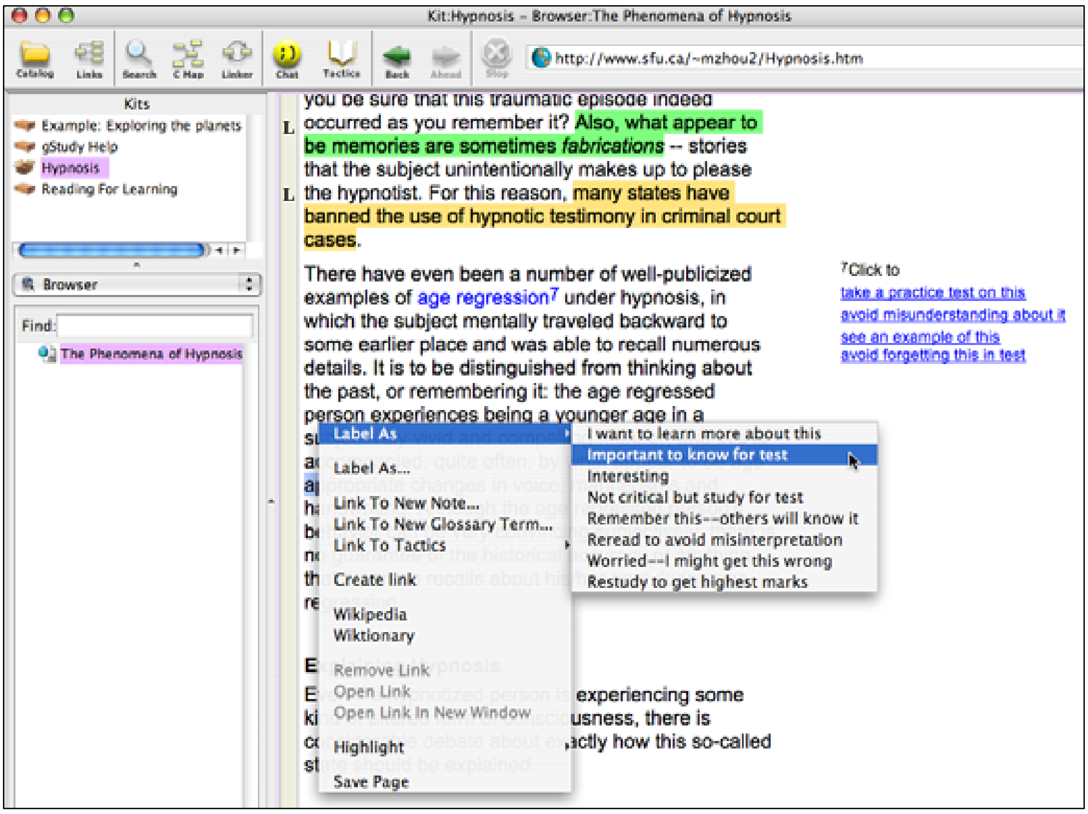


Figure 1. *A screenshot of tagging tools in gStudy (labels and hyperlinks).*

In our approach, the theoretical account of types of goals was exposed to students in the form of tags they could use to catalogue content. As students choose content and tag their selection, they manifest goals they hold. For example, tagging a selection as “I want to know about this” identifies a mastery goal whereas a tag of “Review this to earn higher marks” corresponds to a performance goal. Two students who select the same content can express different goals by the tags they assign to their selection. As well, a student also can apply multiple tags to the same selection of content they study. If this happens, traces provide evidence that the student simultaneously held multiple goals regarding that content.

The key to this method is to justify the labels for the tags so they both (a) represent researchers’ theory of goals as accurately as possible and (b) communicate this message clearly to the student. This was done in two means. First, goal theory provides a basis for composing these terms and phrases. We grounded our work within the currently prevalent 2 x 2 goal framework of goal type and approach/avoidance model developed by Elliot and McGregor (2001). Second, we conducted a pilot study to validate that our interpretation of tags was shared by students.

Data Collection and Analysis

The use of a computer-tracking system provides a number of unique data-collection possibilities that are not available through other conventional techniques. The unobtrusive nature of this method enables researchers to track the experiences of a learner in a non-linear environment without disrupting the learner’s navigation through the program, and more importantly, to obtain data in real-time so that researchers have the facility to replay them together, thus offering a “virtual” re-creation of learners’ actions in that particular situation. As students select and use tools in gStudy, the system writes a detailed log file in XML format, which contains a complete list of time-stamped events tracing learners’ behavior as they navigate through documents, and tag useful information.

To interpret which goals learners establish and how their goal adoptions vary during the engagement in goal-directed activities, it is necessary to capture the context wherein a given type of goal was pursued and how this goal-setting process varies as the contexts change. We developed a parser program called Log Validator [19] to meet these objectives. Figure 2 presents an overview of the analysis. To prepare for parsing a log file with Log Validator, the researcher needs to create an action library that defines each multi-event action in terms of the fine-grained events in the log. Log Validator

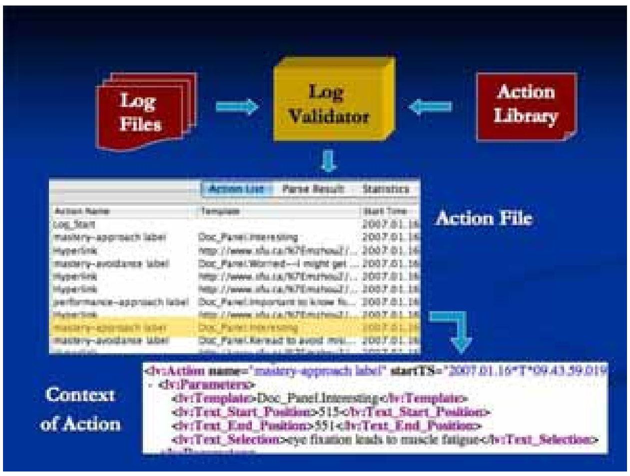


***Figure 2**. Goal-tracing analysis workflow.*

matches events in a log file to the canonical action patterns defined by the researcher in the *action library* and builds an action file consisting of temporally ordered learner actions as well as the context (selected content) to which a goal-oriented label tool was applied could also be extracted. In this way, a fuller picture of how a learner went about learning can be painted. In the example presented in Figure 3, the student adjusted his goal orientation when confronted with different types of contents. Apparently, while shifting frequently between mastery-approach and performance-approach goals,

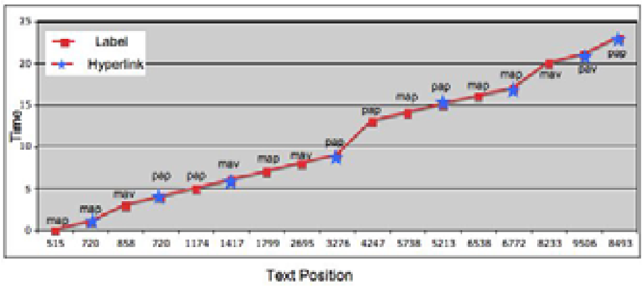


***Figure 3**. Log analysis result example (map = mastery-approach goal orientation; pap = performance-approach goal orientation; mav = mastery-avoidance goal orientation; pav = performance-avoidance goal orientation)*

he also pursued avoidance-oriented goals occasionally, especially on the key concepts whose accurate meanings are more difficult to grasp.

Conclusion

Given reasons outlined above regarding the use of alternative ways to capture the interaction of motivational and cognitive processes beyond conventional surveys or interviews, this state-of-the-art goal-tracing methodology provides researchers with opportunities to address the concerns by testing (a) whether learners’ behavioral choices accord with theoretical accounts of how goal orientation affects studying which determines achievement, (b) whether and how goal orientation develops across contexts and individuals and impacts studying tactic use, and (c) whether different types of goal orientations are adopted within a studying session. To track learners’ conceptual and motivational development as accurately as possible, the trace data should be triangulated using other kinds of data. We believe this “thick description” can provide multiple levels of integrated information about learners’ attitudes, beliefs, and interactions with the content.

References

1. Pintrich, P. R., & Schunk, D. H. (1996). *Motivation in education: Theory, research, and applications*. Englewood Cliffs, NJ: Prentice Hall.
2. Covington, M. V. (2000). Goal theory, motivation, and school achievement: an integrative review. *Annual Review of Psychology,* **51**, 171–200.
3. Pintrich, P. R. (2003). A motivational science perspective on the role of student motivation in learning and teaching contexts. *Journal of Educational Psychology,* **95**, 667–686.
4. Schunk, D. H. (2005). Self-regulated learning: the educational legacy of Paul P. Pintrich. *Educational Psychologist,* **40**, 85–94.
5. Craik, F. I., & Lockhart, R. S. (1972). Levels of processing: a framework for memory research. *Journal of Verbal Learning and Verbal Behavior,* **11**, 684–971.
6. Entwistle, N. (1988). *Styles of learning and teaching*. London: David Fulton.
7. Veermans, M., & Järvelä, S. (2004). Generalized learning goals and situational coping in inquiry learning. *Instructional Science,* **32**, 269–291.

8. Salovaara, H., & Järvelä, S. (2003). Students' strategic actions in computer supported collaborative inquiry. *Learning Environments Research,* **6**, 267–284.
9. Salovaara, H. (2005). *Achievement goals and cognitive learning strategies in dynamic contexts of learning*. Unpublished doctoral dissertation, University of Oulu, Finland.
10. Zhou, M. (2008). *Operationalizing and tracing goal orientation and learning strategy*. Unpublished doctoral dissertation, Simon Fraser University, Canada.
11. Bråten, I. & Samuelstuen, M. S. (2004). Does the influence of reading purpose on reports of strategic text processing depend on students' topic knowledge? *Journal of Educational Psychology,* **96**, 324–336.
12. Järvelä, S., Salonen P., & Lepola, J. (2001). Dynamic assessment as a key to understanding student motivation in a classroom context. In P. Pintrich & M. Maehr (Eds.) *Advances in research on motivation: New directions in measures and methods* (Vol.12, pp.217–240). Greenwich: Elsevier.
13. Järvelä, S., & Salovaara, H. (2004). The interplay of motivational goals and cognitive strategies in a new pedagogical culture. *European Psychologist,* **9**, 232-244.
14. Winne, P. H. & Perry, N. E. (2000). Measuring self-regulated learning. In M. Boekaerts, P. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp.531–566). Orlando, FL: Academic Press.
15. Murtagh, F. E. M., Addington-Hall, J. M., & Higginson, I. J. (2007). The value of cognitive interviewing techniques in palliative care research. *Palliative Medicine,* **21(2)**, 87-93.
16. Tourangeau R. (1984). Cognitive sciences and survey methods. In T. B. Jabine, M. L. Straf, J. M. Tanur, & R. Tourangeau (Eds.), *Cognitive aspects of survey methodology: Building a bridge between disciplines* (pp.73-101). Washington, D. C.: National Academy Press.
17. Perry, N. (2002). Using qualitative methods to enrich understanding of self-regulated learning. Introduction to the special issue. *Educational Psychologist,* **37(1)**, 1–3.
18. Winne, P. H., Nesbit, J. C., Kumar, V., & Hadwin, A. F., Lajoie, S. P., Azevedo, R. A., & Perry, N. E. (2006). Supporting self-regulated learning with gStudy software: The Learning Kit Project. *Technology, Instruction, Cognition and Learning*, **3**, 105-113.
19. Xu, Y., Nesbit, J. C., Zhou, M., & Winne, P. H. (2007). *LogValidator: A tool for identifying and mining patters in gStudy log data*. [computer program]. Simon Fraser University, Burnaby, BC.

A computer method "point-pair system" for measuring movement in facial expression based on video from a natural/clinical context

M. Schiavenato[1] and P. Scovanner[2]
[1]*University of Rochester, School of Nursing, Rochester, USA, martin_schiavenato@urmc.rochester.edu*
[2]*School of Electrical Engineering and Computer Science, University of Central Florida, Orlando, USA, pscovanner@cs.ucf.edu*

Introduction

There are "paper and pencil" facial coding systems in expression research; the FACS, facial action coding system [1], is perhaps the one most widely used. In neonatal pain research, the NFCS, neonatal facial coding system [2], which is derived from the FACS and specific to areas of neonatal pain expression, is frequently used in this field of research. The measurement level derived with this tool is categorical. We attempted to increase the measurement level to ratio with the following "point-pair" methodology.

Point-Pairs

We assigned pairs of points between commonly implicated areas of movement in pain expression; for example, *point-pair 1* lies between the medial borders of the eyebrows; to track horizontal brow movement or equivalent to *brow bulge* in the NFCS vernacular. We devised a total of 7 point-pairs to track movement in the mouth, cheek and eye areas. Whereas previous attempts to track facial action have used video, those attempts assume a stationary head [3], something that was not possible since we were recording infants receiving a painful stimulus (a heel stick), and they almost universally rotated their head in reaction. Also, because these recordings occurred in a naturalistic environment (a newborn nursery), it was not feasible to have additional equipment to facilitate the measurement of video with moving targets. This method assumes that faces are always facing the camera. In order to achieve this, the video was converted to still images (jpegs) using Rad Video Tools (Rad Game Tools, Kirkland, WA). The still pictures were reviewed and two pictures were selected: one depicting a "neutral" face (Figure 1), sometime close to and before receiving the painful stimulus; and another picture depicting a "reaction" face (Figure 2), sometime close to and after receiving the painful stimulus. Thus, because of methodological/practical reasons, we replaced the machine assignment associated with automatic systems, with a manual assignment that tried to maximize the position and expression of the face between the neutral and reaction pictures.

Figure 1. *Neutral face.*

Point-Pair Calculation

The goal was to measure movement, in pixels, in particular facial areas implicated in neonatal pain expression by tracking change between point-pairs. To preclude issues of image size and differences in infant anthropometrics, we chose *percent of facial width* as the standardized unit of measurement. Each child's face-width was measured twice at both baseline and reaction, and the average was used to scale all subsequent point-pair pixel measures. Baseline and reaction images maintained similar scales. Measuring and averaging over both images allowed for a second view and more accurate estimation of facial width. We begin by calculating the Euclidean distance between a pair of points. The final point-point pair output, expressed as a percent of face-width, was the computed distance between the two test points divided by the scale and multiplied by 100:

$$P_i = \frac{100 * dist(X_{i1}, X_{i2})}{dist(S_1, S_2)}$$

Or,

$$P_i = \frac{100 * \sqrt{(x_{i1} - x_{i2})^2 + (y_{i1} - y_{i2})^2}}{\sqrt{(x_{S1} - \mathrm{x}_{S2})^2 + (\mathrm{y}_{S1} - \mathrm{y}_{S2})^2}}$$

Where S are the scale points derived from measuring facial width. Point-pairs were calculated for both baseline and reaction images. Point-pair change, net movement between images, was calculated as:

$$\Delta P_i = P_i^R - P_i^B$$

Where P_i^R is the point-pair value for the reaction image, and P_i^B is the point-pair value for the baseline.

The above calculations were done in Matlab (MathWorks Inc., Natick, MA).

Discussion

The manual assignment of points undoubtedly introduces a level of subjectiveness and some degree of error to the measurement. It should be mentioned that in the earlier phases of this research, we tried an affine transformation to account

Figure 2. *Reaction face.*

for minor changes in head movement; however, these affine transformations were prone to considerable error due to the lack of a good number of rigid points on the face, and again, to considerable head movement. Clearly methodological issues of measuring expression on a moving face remain to be tackled. Our point-pair system demonstrates the ability to advance the measurement of the facial display of pain from a categorical level to a ratio level, thereby making possible statistical comparisons heretofore not afforded with current methodology.

References

1. Ekman P, Friesen WV. (1978). *Facial action coding system*. Palo Alto: Consulting Psychologists Press.
2. Grunau RE, Oberlander T, Holsti L, Whitfield MF. (1998). Bedside application of the Neonatal Facial Coding System in pain assessment of premature neonates. *Pain*, **76**, 277-286.
3. Pantic M, Patras I. (2005). Detecting facial actions and their temporal segments in nearly frontal-view face image sequences. *Presented at IEEE International Conference on Systems, Man and Cybernetics, (Waikoloa, HI. October 10-12, 2005)*.

Polar Coordinate Analysis using Matlab, an application in soccer

Abigail Perea[1], Julen Castellano[2], M. Teresa Anguera[3], and Lorea Alday[1]
[1]*Association for Studies and Research in Sport (GIDE), Vitoria, Spain, bigaper@hotmail.com*
[2]*Department of Sports, University of the Basque Country, Vitoria, Spain*
[3]*Department of Behavioral Sciences Methodology, University of Barcelona, Spain*

The polar coordinate analysis [1] is a double data reduction strategy which provides a vector representation which determines the relation between the criterion behavior, this is, between the criterion which is our object of study and the other categories which conform the taxonomic system. With the intention of automatically perform the polar coordinates analysis a Matlab [2] script file was created.

Six matches played during the final phase of the Germany '06 World Championship were selected for this study, in which a taxonomic system (ad hoc) was defined with the objective of specifying the interaction contexts performed by the teams and the use of the space during the game. The observation and data registering was performed using SOCCAF v2.2video [3]. Afterwards, a sequential analysis using SDIS-GSEQ [4] and a polar coordinate analysis were performed.

The results obtained from the polar coordinates analysis allow describing the game action in soccer, adding the diachronic dimension of the events and combining the prospective and retrospective perspectives. By automatically performing the calculations using the Matlab approach, we have dramatically reduced the time for the calculations, as well as increased the quality of the results.

References

1. Sackett, G. P. (1980). Lag Sequential Analysis as a data Reduction Technique in Social Interaction Research. In D. B. Sawin, R. C. Hawkins, L. O. Walker y J. H. Penticuff (Eds.), *Exceptional infant. Psychosocial risks in infant-environment transactions*. New York: Brunner/Mazel, pp. 300-340.
2. Matlab. (1994-2005).The Math Works, Inc.
3. Perea Rodríguez, A. E., Alday Ezpeleta, L. and Castellano Paulis, J. (2006). Aplicación informática específica para la observación de la acción de juego en fútbol. In *Castellano Paulis, J.; Sautu Apellaniz, L. M.; Blanco Villaseñor, A.; Hernández Mendo, A.; Goñi Grandmontagne, A. y Martínez de Ilarduya, F. (coord.), Evaluación e intervención en el ámbito deportivo*. Diputación Foral de Álava, pp. 285-294.
4. Bakeman, R. and Quera, V. (1996). Análisis de la interacción. *Análisis secuencial con SDIS y GSEQ*. Madrid: RA-MA.

GamE: *G*ames *a*s a *m*ethod for *e*liciting *e*motions

Suleman Shahid, Emiel Krahmer, and Marc Swerts
Department of Information and Communication Sciences,University of Tilburg, The Netherlands
{S.Shahid, E.J.Krahmer, M.G.J.Swerts}@uvt.nl

In this paper, we discuss the design of a simple but effective card game developed under the 'GamE (Game as a Method for eliciting emotions) paradigm'. Under the GamE paradigm, the objective is to design a number of games to be used as a tool for inducing emotions in a natural and ethical way and this card game is one example of this.

Research on emotions has grown massively in the past few years which resulted in new theories, methodology and interesting findings. This research has been done from different perspectives (evolutionary, psycho-physiological, neurological, psychodynamic, etc) where every perspective took a meticulous approach in understanding and conducting research about emotions. Regardless of the different point of views, a fundamental focus in all perspectives has always been on developing innovative methods for inducing emotions in a natural and ethical way, which is still a challenging task [1]. There are a number of methods used for inducing natural emotions but usually the results gained from these techniques are not easy to generalize because of their artificially controlled settings [2] and dependability on a particular culture. There is always a need of developing innovative techniques, which could not only be used across cultures but also in a natural environment.

One technique, which we explore here, is to use games as emotion inducers. It is well known that players from different age groups can become emotional (both negatively and positively) while playing games and games give them a very engaging experience [3]. Based on this general idea, we developed a simple card game to use as a tool for inducing emotions in children. The card game is developed using Microsoft® PowerPoint®, and every game consists of six cards. When the games starts, only the first card is visible ('3' in the case of figure 1) and the other five cards are placed upside down so the numbers are hidden. Numbers on all cards are between 1 to 10 and a particular number cannot repeat in a single game (i.e. in the above example the number '8' will not repeat in next 5 cards). Appropriate colourful images are chosen for the game background and different animations are used to turn card around for making the game more attractive for children.

Figure 1. *Winning variant of the game*

When a game starts, the task of the player(s) is to guess whether the upcoming number on the next card will be lower or higher than the number ob the previous card. Once players have made their guess, the relevant card is turned around and the number on the card is made visible on the screen. In addition to this, a characteristic non-speech audio sound is also played right after the card is visible which inform players about the correctness or incorrectness of their answer. If players guess correctly then they are asked to guess the next number and they only win the game if they guess all the cards in a game correctly. If players predict the number incorrectly at any stage of the game, they immediately lose the game and move to the next game.

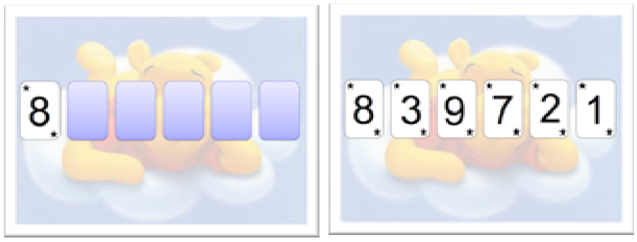


Figure 2. *Loosing variant of the game*

An important design consideration of the game is its deterministic nature. So far, we have used sequences of cards, where players will win half of the time and lose half of the time, if they make "rational choices". Figure 1 is an example of the winning variant where guessing the sequence of 'bigger – smaller – bigger…' will lead to a win. Furthermore in figure 1, at card 5, the rational choice will be a 'lower' number because no number can be greater than 10. Similarly for the last card, the choice will be 'lower' because the number higher than 9 has already been displayed at card 4. Figure 2 represents a losing variant where the most probable outcome for the final card would be a number higher than 2, but guessing "higher" would make this a losing game. Winning and losing games are typically mixed in the sequence, starting and ending with a variant in which children were likely to win.

In a series of experiment, we have used this card game to collect emotional responses to winning or losing the game. So far, we collected data from 144 children (half 8 years old and half 12 years old) who played the game either in the Netherlands or in Pakistan, and either individually or in pairs. Figures 3 shows the representative stills of winning and losing of Dutch and Pakistani children belonging to both age groups. Generally, the game worked quite well and as intended. Almost all individual and pairs of participants indeed made the logical choices that were expected in most of the cases, so that each individual child and pair of children lost at least two games and won at least two games. Additionally, not even a single child noticed and reported that the game was in fact a deterministic simulation. Rich data was which so far has been used in number of perception tests where viewers from Pakistan and the Netherlands watched and judged the emotional response of Pakistani and Dutch children when they win or lose a game. The details of these studies will be described elsewhere.

Figure 3*. Representative still for (Left to Right): Pakistani girl loosing, Dutch couple winning, Pakistani couple winning and Dutch boy loosing.*

On the basis of these early results, we can conclude that 'basic' games can be a useful tool for inducing positive and negative emotions naturally and ethically. Furthermore, this is a kind of emotion induction method that goes beyond traditional experimental settings and gives valuable insights about emotion regulation in natural settings. In the future, we would like to run more experiments with the same game in different cultures and would like to develop more extensive (but still simple) games under the same GamE paradigm for using them as an emotion inducer.

References

1. Cornelius, R.R. (1996). The science of emotion. research and tradition in the psychology of emotion. Upper Saddle River, NJ: Prentice-Hall.
2. Kaiser, S. & Wehrle, T. (1996). Situated emotional problem solving in interactive computer games. In N. H. Frijda (Ed.), *Proceedings of the IXth Conference of the International Society for Research on Emotions* (pp. 276-280). Toronto: ISRE Publications.
3. Lazzaro, N. (2004). Why We Play Games: Four Keys to More Emotion Without Story. In Game Developer Conference (GDC), san Fransisco, USA.

Toward music structure of improvised drum performances

Rumi Hiraga[1], Noriyuki Matsuda[2], and Nobuko Kato[1]
[1]*Faculty of Industrial Technology, Tsukuba University of Technology, Japan, rhiraga@a.tsukuba-tech.ac.jp*
[2]*Graduate School of Systems and Information Engineering, University of Tsukuba, Japan*

Our purpose is to find out the methods of utilizing acoustic features of improvised drum performances expressing an intended emotion. Since we plan to build a system that assists hearing-impaired people in playing a drum set and making music communication by sharing an emotion, we have conducted a series of cognitive experiments to investigate the communication possibility of improvising drum performance with an intended emotion. Understanding improvised drum performances from the point of view of acoustic features is also of concern in building the system. In this paper, we describe a method to find out music structure of improvised drum performances as the first step to utilize the acoustic features in our system.

There have been several works in music information retrieval such as to detect chorus [1] and segmentation labeling [2, 3]. Some studies work on understanding emotion in music from acoustic features [4, 5]. One of the differences of our work from others is that our music pieces are played in an improvised style. It signifies that there is not a strict structure in our music performance, on the other hand classic and popular music can be called well-structured. Our music has freedom in terms of beat, measures, and repeating structures; these structural components can expand and shrink at the discretion of a player and there is not an assumable strict music structure consisting of music components of verse, chorus, and middle eight, for example. In order to obtain the methods of utilizing acoustic features that relate to emotion, we started from investigating what we can find with acoustic features of our music.

We have a set of drum performances that a professional drummer played various percussion instruments in an improvised style. He played a bass drum, a snare, a set of concert tom, a Chinese gong, and a suspended cymbal using drumsticks, mallets, brushes, and other items. He played a few sets of performances on these instruments, with each set consisting of four performances representing one of the four emotions (joy, fear, anger, and sadness).

We used MIR Toolbox [6] for MATLAB to divide a piece into segments and get musical features of each segment. The segmentation method is based on media segmentation [7] that uses spectrum and onset peaks to get the correlation between frames. The detected segments were the music components as verse or chorus for a well-structured music piece. Because our music is played in an improvised style only with the percussion instruments, the detected segments were much shorter than those in well-structured music. Thus, a so-called music component is consisted of a few or more of the consecutive segments in our music.

Then we made a similarity matrix by calculating the distance between segments using ten acoustic features (number of onset in a segment, centroid, brightness, spread, skewness, kurtosis, roll off, flatness, roughness, and irregularity). It is followed by two methods for observing the relationships among segments: (1) Draw an undirected graph using Graphviz [8] where nodes are segments and edges represent the closer distance between segments. Empirically we defined the threshold for an edge between a segment S1 and S2 to appear as follows: An edge is drawn between S1 and S2 if the distance between these two segments is smaller than m-2σ where m is the mean value of the distance of two segments in a music piece and σ is its standard deviation. In this way, we can find isolated segments. (2) Compute Singular Value Decomposition (SVD) of the similarity matrix, get unit norm matrices scaled by singular values, and visualize the norm matrices.

The finding was that isolated segments in an undirected graph can be the boundary of a music component; usually they are the beginning of the component. Unit norm matrices with the large singular values also show the possible boundaries that are consistent with the isolated segments obtained in the graph. Furthermore, when we clustered the isolated segments with k-means clustering where k=2, then segments in a cluster with the smaller number of segments are peculiar ones in the music, such as the music-dividing segment. We can find those segments usually in the unit norm matrix with the largest singular value. This way of finding music structure may help people understanding a musical piece that have no strict musical structure.

References

1. Goto, M. (2003). A chorus-section detecting method for musical audio signals, Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing 2003, V-437-440.
2. Aucouturier, J.-J. and Sandler. M. (2002). Finding repeating patterns in acoustic musical signals: applications for audio thumb nailing. *AES 22nd International Conference on Virtual, Synthetic and Entertainment Audio*.
3. Chu, S. and B. Logan (2000). Music Summary Using Key Phrases. Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing.
4. Bresin, R. and Friberg, A. (2000). Emotional coloring of computer-controlled music performances, 24:4, *Computer Music Journal,* MIT Press, 44-63.
5. Schubert, E. and Fabian, D. (2005). An experimental investigation of emotional character portrayed by piano versus harpsichord performances of a J.S. Bach excerpt, *Mackinlay. E. (Ed), Aesthetics and Experience in Music Performance.* Cambridge Scholars Press, 7-94.
6. Lartillot, O. and Toiviainen, P. (2007). MIR in MATLAB (II): A Toolbox for Musical Feature Extraction from Audio. *Proceedings of International Conference on Music Information Retrieval.*
7. Foote, J. T. and Cooper M. L. (2003). Media Segmentation using Self-Similarity Decomposition. *Proceedings of SPIE Storage and Retrieval for Multimedia Databases*, *Vol.* ***5021****, (San Jose, CA, January 2003)*, 167–175.
8. Graphviz. http://www.graphviz.org/

Measuring subjective well-being of persons with profound intellectual and multiple disabilities

S. Dero, K. Petry, and B. Maes

Faculty of Psychology and Educational Sciences, Centre for Parenting, Child Welfare and Disabilities, Catholic University of Leuven, Leuven, Belgium, Sarah.Dero@ped.kuleuven.be

One of the most difficult challenges in the quality of life-research is the measurement of the subjective well-being of people with profound intellectual and multiple disabilities (PIMD). This group is characterized by profound cognitive disabilities (IQ<20) and severe motor and/ or sensory limitations and medical problems, often caused by severe brain damage. Since persons with PIMD lack the necessary cognitive and communicative abilities to make a global judgment of their life, self-reports on their subjective well-being are not applicable. A frequently used alternative is the 'proxy-approach', in which someone who knows the person well forms an opinion about several life aspects instead of the person him/ herself. The literature however yields conflicting results concerning the psychometric value of this proxy-approach [1].

Research on the subjective well-being of persons with PIMD focuses increasingly on the 'hedonic level' of the subjective well-being – this is defined as 'the individual's expressions of positive or negative emotions or moods'. Direct emotional reactions are complex and multifaceted. Measuring the hedonic level of persons with PIMD, therefore, needs to focus on the behavioral, nonverbal and/ or physiological aspects. Recently some procedures have been developed that imply an observation of behavioral and nonverbal expressions of positive and negative affect [2, 3, 4]. Although these procedures look promising, they still need further validation. To date, the physiological component of emotional reactions has not been included in measures of the subjective well-being of persons with PIMD. In research on emotions some physiological variables have been successfully used to measure emotional responses [5]; they also seem to distinguish between positive and negative emotions and give information about the frequency, the intensity and, to a minor extent, the variability of emotions. Next to this, this approach might offer a possibility to evaluate emotions in a more objective way and without relying on the intermediary position of a proxy or a researcher.

Since physiological parameters are not yet considered as indicators of positive or negative emotions on the subjective well-being of persons with PIMD, a preliminary investigation is required. The aim is to validate the measurement of physiological parameters of emotional reactions. To gain a sharpened understanding of this approach, we will set up a case-study research. Four adult persons with PIMD (over 18 years of age) will be selected to participate in the study. In several daily situations – corresponding to the direct behavioral procedure – 'extreme' expressions of positive and negative emotions/ mood will be manipulated. Because of ethical reasons and the enormous heterogeneity of people with PIMD, members of direct support staff will be involved in defining the stimuli for each participant. During these situations several physiological parameters will be registered continuously. We selected parameters which have been used previously in research on subjective well-being and are usable in the target group, namely heart rate, heart rate variability, bodily temperature, respiratory function and skin conductance. The physiological parameters will be measured by means of easy-to-use non-invasive ambulatory technology. At the same time the situations will be recorded on videotape. In this way, it might provide hypotheses which may be tested systematically with a larger number of cases in the main study of this research project. Next to investigate the validity of this methodology, the application of the ambulatory technology to measure the physiological parameters will be examined. Finally, the selection of physiological measuring will be refined or enlarged.

In the main study the aim is to examine a set of measures to assess the 'hedonic level' of subjective well-being of persons with PIMD. To achieve this aim, 25 people with PIMD will be selected. Three measures will be administered for each person with PIMD in the sample within a four week period.

(1) In week one, the direct behavioral procedure will be set up. On the basis of an interview with a member of direct support staff and a parent/ family member, an affective communicative profile of the person with PIMD will be drafted. At the start of week two, several daily situations will be recorded on digital videotape during a two-week period. Finally, these video recordings will be analyzed. The researcher codes the frequency and the intensity of expressions of positive and negative affect, using the earlier drafted affective profile. In order to extract observational data from the video recordings, professional software will be used. Based on these data, the frequency, intensity and valence of affect will be calculated.

(2) During the video recordings of the selected daily situations, several physiological parameters will be registered continuously. The selection of the phsyiological parameters and the non-invasive ambulatory technology is based on the preliminary investigation.

(3) In week four, two members of the direct support staff will complete an informant questionnaire, namely the MIPQ (Mood, Interest & Pleasure Questionnaire). The MIPQ is an indirect observational measure that addresses the nonverbal as well as the behavioral component. The informants will be asked to rate operationally defined observable behaviors, which relate to the construct of mood, interest and pleasure, on a five-point Likert-type rating scale. The MIPQ gives information on the valence and the frequency of affect.

To answer our research questions, several relations will be explored, for instance using correlation coefficients: between the frequency data of all three measures, between the intensity data of the physiological measures and the direct observational procedure, between the frequency, intensity or variability data for the positive and negative emotions, between the indirect observational and direct observational measures, etcetera. For the MIPQ, reliability will be determined on the basis of internal consistency and inter-rater reliability. For the direct observational procedure, 25% of the material will be double coded by a second researcher.

References

1. Perry, J., Felce, D. (2002). Subjective and objective quality of life assessment: Responsiveness, response bias, and resident: proxy concordance. *Mental Retardation,* **40**, 445-456.
2. Lyons, G. (2005). The Life Satisfaction Matrix: an instrument and procedure for assessing the subjective quality of life of

individuals with profound multiple disabilities. *Journal of Intellectual Disability Research,* **49(10)***,* 766-769.

3. Petry, K., Maes, B. (2006). Identifying expressions of pleasure and displeasure by persons with profound multiple disabilities. *Journal of Intellectual and Developmental Disability,* **31(1)**, 28-38.

4. Ross, E., Oliver, C. (2003). The assessment of mood in adults who have severe or profound mental retardation. *Clinical Psychology Review,* **23***,* 225-245.

5. Cacioppo, J.T., Berntson, G.G., Larsen, J.T., Poehlmann, K.M., Ito, T.A. (2000). The psychophysiology of emotion. In M. Lewis & J.M. Haviland-Jones (Eds.), *Handbook of Emotions, 2nd Ed. (pp. 173-191). New York: Guilford.*

Facial Expression recognition based on dynamic textures

J. Trojanová and M. Železný
Department of Cybernetics, Faculty of Applied Sciences, University of West Bohemia, Czech Republic
trojana@kky.zcu.cz, zelezny@kky.zcu.cz

In this paper, we introduced a novel method for facial expression recognition in video scenes based on dynamic textures approach. The dynamic textures are presented as linear dynamic system that models spatial and temporal stochastic process. The dynamic textures are used for learning and recognition of facial expression. Originally dynamic textures were used for video scenes that exhibit certain stationarity properties in time (e.g. Sea-waves, smoke) [1].

Overview on dynamic textures

Brief overview is presented here, more details about the dynamic textures can be found in [1]. The sequence of images is referred as $\{I(t)\}_{t=1..\tau}$, where τ is time interval. The output of the system $y(t) \in R^m$, $t=1,..,\tau$ is used for representing the pixel intensity. The system is defined as follows :

$$x(t+1) = Ax(t) + v(t) \qquad x(0) = x_0;\ v(t) \sim N(0,Q)$$
$$y(t) = Cx(t) + w(t) \qquad w(t)\ N(0,R)$$

where $x(t) \in R^n$ is hidden state, $A \in R^{nxn}$ is the state transition matrix that represent the dynamics of the system, $C \in R^{mxn}$ is the output matrix, v(t) is the driving output to the system, which is assumed to be Gaussian white noise, w(t) is measurement noise.

Estimation of the dynamic textures parameters is done by identification of the system. For the sequence of τ frames the matrix $Y_1^t = [y(1),..,y(\tau)] \in R^{mx\tau}$ is defined. To obtain a unique solution of the system two assumptions about the model is made: canonical model of matrix C: $C^TC=I_n$ (I_n is the *nxn* identity matrix) and data dimensionality reduction $m >> n$, $rank(C)=n$. The estimation of the parameters of the models are:

$$\hat{A} = \Sigma V^T \begin{bmatrix} 0 & 0 \\ I_{\tau-1} & 0 \end{bmatrix} V \left(V^T \begin{bmatrix} I_{\tau-1} & 0 \\ 0 & 0 \end{bmatrix} V\right)^{-1} \Sigma^{-1}; \qquad \hat{C} = U; \qquad X = \Sigma V^T$$

where matrix *U* and *V* are obtained by singular value decomposition of.

$$Y_1^t = U \Sigma V^T$$

Matrix $U \in R^{mxn}$, $U^TU=I$; $V \in R^{\tau xn}$, $V^TV=I$.

Preliminary testing

Each dynamic texture is characterized by parameters: matrix *A*, *C*. To compare textures difference, distance matrix using subspace angle is used. Comparison is as follows. Resolution of the video is 160x120, each frame is divided into 64 blocks. We concentrate on block series 27-30 that represent upper face.

The comparison between neutral state to facial action unit AU02 (rise eye borrows) was made, see Figure 1. Action unit is a term of facial muscle movement defined by Ekman in [2].

Future work

Since dynamic texture capture a dynamic process we assume that training on small number of speakers would reach efficient results even for speakers that were not included in training phase. The testing on larger dataset was not done yet. The database UWB-07-EFER (UWB stands for University of West Bohemia, 07 for year of recording, EFER for Emotions and Facial Expressions Recognition) was recently acquired. The video database consist of facial action units and expressions of 20 people, it is now being annotated.

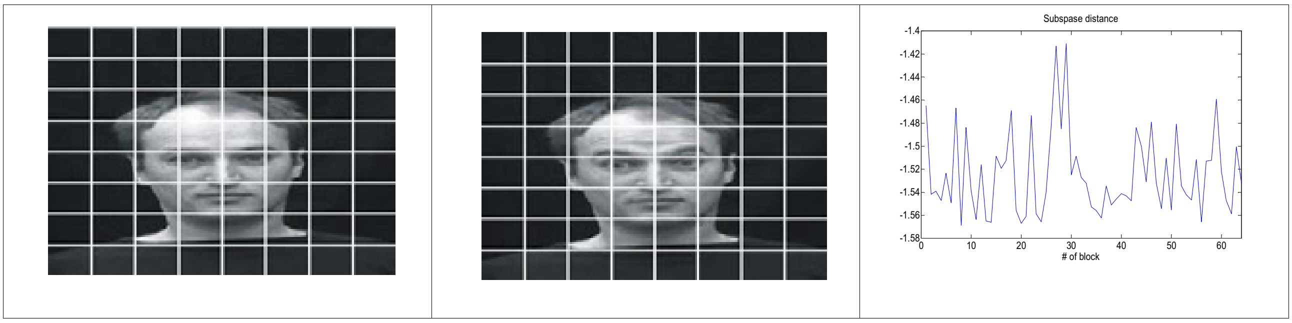


Figure 1. *Comarison between neutral face sequence and sequence that represent eye borrows rise. Image is split into 64 blocks. Plot shows the subspace distance between parameters A,C for all blocks. The biggest distance can be seen on blocks segments 27-30.*

Conclusion

The algorithm is efficient to learn the dynamics of the facial expression. The parameters of the model that represent data can be used as parameterization of the data. The method has shown promising results on preliminary testing.

References

1. Dorretto, G. A. Chiuso, Y.N. Wu and S. Soatto. (2003). Dynamic Textures, *International Journal of Computer Vision*, **51(2)**, 91-109.
2. Ekman P. and Friesen W. (1978) *Facial Action Coding System*, Consulting Psychologists Press Inc. Palo Alto, CA.

Children's interactive behavior with exhibits in an informal educational environment

S. F. Chen[1] and S. Y. Chien[2]
Department of Early Childhood Education, National Taitung University, Taiwan, R.O.C.
[1]shufangnttu@gmail.com, [2]linyu8888@nttu.edu.tw

The purpose of this study is to explore the play and exploratory behaviors of children in an experimental setting which was designed as an informal educational environment. In order to foster a creative learning experience of children in remote area in Taiwan, the research team created a series of self-directed instructional materials under the theme of "the Wonderful World of Insect." An informal educational exhibition was set up and children from the age of 4 to 15 years old were invited to attend the exhibition. Children were free to explore and play with any materials at their own will and at their pace. In order to understand how the instructional materials in informal educational environment stimulate children's behaviors as well as how the child use and learn through these materials, two video machines were set up and children's interactive behaviors with the materials were tapped. In total 164 behavior segments were collected and analyzed. We employed The Observer® XT as data analysis technique and used Spss12.0 for statistics analysis. The results show that a significant difference exits among different types of instructional materials. The patterns of children's interactive behaviors will be described and explained in this paper. Suggestions for designing better instructional materials for the informal educational environment will be made. The techniques of using The Observer® XT to assist the analysis of video data will also be discussed.

Keywords

Instructional materials, informal education, children's play, interactive behavior

References

1. Gazit, E., & Chen, D. (2003) Using The Observer to analyze learning in virtual worlds. *Behavior Research Methods, Instruments, & Computers,* **35(3)**, 400-407.
2. Kola-Olusanya, A. (2005) Free-choice environmental education: understanding where children learn outside of school. *Environmental Education Research,* **11(3)**, 297–307.
3. Mayfield, M. I. (2005) Children's museums: purposes, practices and play? *Early Child Development and Care*, **175(2)**, 179–192.
4. Rix, C., & Mcsorley, J. (1999) An investigation into the role that school-based interactive science centres may play in the education of primary-aged children. *International Journal of Science Education*, **21(6)**, 577-593.
5. Speaker, K. M. (2001) Interactive exhibit theory: hints for implementing learner-centered activities in elementary classrooms. *Education*, **121(3)**, 610-614. (AN 4434387)
6. Tunnicliffe, S. D., & Laterveer-de Beer, M. (2002) An Interactive Exhibition about Animal Skeletons: Did the Visitors Learn Any Zoology? *Journal of Biological Education*, **36(3)**, 130-134. (EJ649591)
7. Semper, R. J. (1990) Science museums as environments for learning. *Physics Today*, **43(11)**, 50-56.
8. Werby, O. (1994) The relationship between changes in perceptual focus and understanding. Paper presented at the Annual Meeting of American Educational Research Association. (ED370971)
9. Bowker, R. (2002) Evaluating Teaching and Learning Strategies at the Eden Project. *Evaluation and Research in Education*, **16(3)**, 123-135. (EJ789479)
10. Boram, R., & Marek, E. A. (1991) *The Effects of Free Exploration from Hands-On Science Center Exhibits*. Paper presented at the Annual Meeting of the National Association for Research in Science Teaching. (ED337354)
11. Noldus, L. P. J. J., Trienes, R. J. H., Hendriksen, A. H. M., Jansen,H.& Jansen, R. G. (2000) The Observer Video-Pro: new software for the collection, management, and presentation of time-structured data from videotapes and digital media files. *Behavior Research Methods, Instruments and Computers*, **32**, 197-206.
12. Visser, M. E. (1993) The Observer, a software package for behavioural observations. *Animal Behaviour*, **45**, 1045-1048.

Observed toy preference as a measure of gender behavior in an epidemiological study examining the neurotoxicological effects of environmental exposure

G. Vermeir[1,2], R. Verachtert[1], M. Thijs[1], and M.K. Viaene[1,3]
[1]*Flemish Agency for Mental Health (OPZ) Geel, Belgium, griet.vermeir@opzgeel.be*
[2]*Faculty of Medicine and Health Sciences, Ghent University, Belgium*
[3]*Faculty of Medicine, Kuleuven, Belgium*

Introduction

Environmental exposures to a variety of neurotoxins remain a concern in western industrialized countries, as well as in developing countries, especially concerning health problems in vulnerable populations such as the unborn child. PCB's and dioxine-like compounds are regarded to belong to the most ubiquitous developmental neurotoxins [1].

As PCBs and dioxine-like compounds have an endocrine disrupting effect, it is interesting to notice that play behavior is reported to change in boys and girls (less masculine and more feminine play behavior) according to the level of prenatal PCB exposure. Some data in girls suggest that even the postnatal PCB exposure might influence play behavior, although in the opposite direction (more masculine) [2]. Although these data are certainly not conclusive, as the questionnaire (PSAI) which was used was not appropriate for the age of the children and test-retest results were rather poor (0.6) [3]

In addition, this questionnaire was filled in by the mother and is therefore an indirect measure of the child's behavior. To replicate these findings, we used an alternative and more direct method to measure gender behavior in children.

Method

The present study is a part of the Environmental Health Action Program (EHAP) (2002-2006) in Flanders, investigating environmental exposure to e.g. Pb, Cd, PCBs, compounds with 'dioxin-like' , chlorinated pesticides [DDE and hexachlorobenzene (HCB)], and organic solvents in eight selected regions of Flanders. In this part of the EHAP, the relation between neuro-developmental parameters (e.g. IQ, milestones, language development) and the prenatal exposure markers of PCB's and dioxin-like compounds will be studied. Measuring development in the different domains of human brain function is a lengthy and difficult task. Therefore an additional aim is to investigate if behavior (questionnaires, play behavior observations) is a useful bio-effect parameter in follow-up studies in populations exposed to toxic agents.

"Toy preference" is assessed by measuring the time a child is playing with toys which are generally preferred either by girls (feminine) or by boys (masculine), or which is liked by both equally (neutral) [4]. The toys consisted of four feminine toys (a babylike doll, two barbie dolls, a thee set and a hairstyling doll), two neutral toys (a book and a puzzle), and four masculine toys (a firetruck, four little cars, building blocks and a gun). The toys were presented in a standardised way. The objects were arranged in a semi circle around the child in the following order (clockwise): gun, thee set, firetruck, two barbie dolls, puzzle, book, four little cars, hairstyling doll, building blocks and babylike doll. All objects were visible and within reach for the child. A 7 minutes play was videotaped. The total time playing with feminine, masculine and neutral toys respectively and the time spent without playing was calculated.

This task was reported to have a high test-retest reliability [5], which was our own experience (5 videotapes were scored by two different persons and the difference was maximum some seconds per toy and consequently less then ten seconds per category, not changing anything to the percentual preference).

Results

At this moment, the data-analyses are still ongoing. The first preliminary results (N=70) suggested that higher exposed boys played less with boy specific toys.($p<.05$) and more with gender non-specific toys ($p<.01$).

References

1. Guo YL, Lambert GH, Hsu CC, Hsu MML (2004). Yucheng: health effects of prenatal exposure to polychlorinated biphenyls and dibenzofurans. *International Archives of Occupational and Environmental Health,* **77,** 153-158.
2. Vreugdenhil HJI, Slijper FME, Mulder PGH, Weisglas-Kuperus N. (2002). Effects of perinatal exposure to PCBs and dioxins on play behavior in Dutch children at school age. *Environmental Health Perspectives*, **110(10)**, 593-598.
3. Kaufman, AS (2003). Critique of Vreugdenhil et Al.'s Study Linking PCBs to the Play Behaviors of Dutch Girls and Boys. Environmental Health Perspectives, **111**.
4. Connor, J.M., & Serbin, L.A. (1977). Behaviorally based masculine- and feminine-activity preference scales for preschoolers: Correlates with other classroom behaviors and cognitive tests. *Child Development*, **48**, 1411–1416.
5. Henserson, BA, Berenbaum, SA (1997). Sex-typed play in opposite-sex twins. *Developmental Psychology*, **31(2)**, 115-23.

Analysis of children behavior in space for children through User Observer method

Kyung Ran Choi[1] and Yuan Huang[2]
Graduate School of Techno Design, Kookmin University, Seoul, South Korea,
[1]ran@kookmin.ac.kr, [2]ruomu@naver.com

Abstract

User Centered Design (UCD) is a design philosophy with primary focus on understanding the user and the user experiences to understand the needs of the user. User Research is a part of the UCD process that is used to pinpoint the needs of the user. At this stage, the most important factor to be considered in creating effective overall design is the cultural element.

Especially in the case of space for children, understating various cultural backgrounds that affect factors like, physical changes, psychology, emotion, intelligence and communications for children is paramount. It must be noted that, if only existing research data is used during this stage, some of the details necessary for complete understanding may be overlooked. In order for a complete analysis on desires of users to be conducted, understanding and interpretation of cultural backgrounds through multifaceted analysis is necessary.

User Observation was selected as the research method for cultural element research analysis. We set out to understand the behavior of Korean and Chinese children in a given space by using methods that improved existing typical and quantitative analysis through systematic analytical framework on cultural elements. Children aged 4-7 were selected as the subject for the study. Kindergartens and other spaces built for children in Korea, and kindergartens in China were selected for the sites of the study.

Still photography and Video Ethnography were methods used to observe the children. The data was collected in places of education, play and rest, and The Observer program by Noldus IT was used to organize them by behavior. Cultural elements were the main focus of the analysis. User analysis was separated into three comparisons, relationships between: user and space, user and object, and object and space. Within these three comparisons, cultural elements that occur in each of the three relationships were deduced to separate and identify specific behaviors of children

This study uses User Observation method to recognize and analyze cultural behaviors. The focus of the research is to confirm the viability of using behavioral data in user need based design proposals for space, products and furniture built for children. This study also finds meaning in applying cultural elements based analysis to design planning processes to increase the user-centric design values.

References

1. The implementation of cultural human factors from chosun Dynasty, Kyung Ran Choi, IDEC (Interior Design Edncater Council), Chicago, U.S.A
2. <Design & Cultural Identity>,Kyung Ran Choi, Furniture Design Planning Center, Design House, 1999.
3. Cultural Experience,Ethnograghy in Complex Society, James P, Spradley & David W. McCurdy, prospect, Waveland Press,Inc., 1988.
4. Spatial Behavior, Reginald G. Golledge, Robert J. Stimson,The Guilford press, New York, 1997.
5. User Centered Design, Yamazaki Kazuko, Published by Sea zong, 2005.
6. Real World Research, Colin Robson, Blackwell, Oxford UK & Cambridge, U.S.A., 1996.
7. What to Do with a Human Factor, Rick E. Robinson, American centre for design, 1993
8. Design Research: Methods and Perspectives, by Brenda Laurel ,October 1, 2003
9. Observing the User Experience: A Practitioner's Guide to User Research
10. Mike Kuniavsky| Morgan Kaufmann Publishers.
11. Cultural Analysis, Robert Wuthnow, James Davison Hunter, Albert Bergesen, Edith Kurzweil, 1984, Routledge & Kegan Paul plc.
12. The Design of everyday things, Donald A. Norman, 1990.
13. Interior Environment Psychology behavior , Lee Youn Suk, Published by Youn Sea Univ. 1998.
14. Environment Psychology and Human Behavior, Lim Seung Bin, Published by Bo Mun Dang, 2007.
15. Investigation Research methodology, Lee Won Wook, Published by Hyung Sul, 2000.
16. Design and Human Psychology, Donald A.Norman, Lee Chang woo. Kim Young jin. Park Hang Ho, Published by Hak ji sa
17. Modern culture theory, Inoue Shun, Published by Lee Hwa Wome`S Univ., 2004.
18. Space Culture political science, Lee Mu Yong , Published by Non Hyung, 2005.

Seeing the complete picture: Combination of telemetric monitoring and behavioral monitoring Special Interest Group

Klaas Kramer
Vrije Universiteit, Amsterdam, The Netherlands, K.Kramer@dienst.vu.nl

This SIG is intended to bring together researchers interested in the combination of physiological data obtained via implantable radio-transmitters and obtained from behavioral observation tools.

Leading experienced researchers will share the fruits of their years of laboratory experience working with fully implantable laboratory telemetry. Their presentations will communicate the details of the telemetry applications and methods they have employed to dramatically reduce animal use, to improve data quality, to do better science, to maximize the amount of information gained from each experimental animal, and to reduce animal stress and discomfort. An emphasis will be placed on the benefits of combining behavioral data obtained by different methods with physiological data.

Special Interest Group Contents

Transportation as major life-event in rats: Effects on welfare and limits of adaptation

J.W.M. Arts[1,2], F.Ohl[1], and K. Kramer[1,3]

[1] Department Animal, Science and Society, Division Laboratory Animal Science, Utrecht University, Utrecht, The Netherlands, j.arts@uu.nl

[2] Harlan NL, Horst, The Netherlands

[3] Department Health, Safety and Envronment, Free University, Amsterdam,The Netherlands

Transportation of laboratory rodents

Transportation is a major stressor in the life of a laboratory rodent. Nevertheless, very little is known about the size of the effect of transportation on the animal and how long it takes for the animal to restore. At the same time, the effects of a stressor on research-results can be radical. Most experiments make use of an acclimatization-period after transport to decrease the influence of transportation on results, but the duration of this acclimatization-period is scarcely based on scientific research. This research project aims on the physiological and ethological effects of transportation on small laboratory rodents.

Article 5 of Appendix A of the European Convention for the Protection of Vertebrate Animals used for Experimental and Other Scientific Purposes (ETS no 123) states *'….even if the animals are seen to be in sound health it is good husbandry for them to undergo a period of acclimatization before being used in a procedure. The time required depends on several factors, such as the stress to which the animals have been subjected which in turn depends on several factors such as the duration of the transportation and the age of the animal. This time shall be decided by a competent person'*.

The question that arises is that even if one is confident about the health status of incoming animals, one still has to decide how long animals should be maintained in the quarantine area so that they can fully acclimatize. In general, animals subjected to the environmental changes occurring during transportation (housing in transport boxes, several hours of travel, final placement in a new facility, exposure to new animal caretakers and procedures) react with changes in their physiology, such as body weight, plasma hormonal levels, heart rate and blood pressure changes [2,5,7-11].

To foster good scientific practice, animals should be used only in experimental procedures after adaptation to their new situation and stabilization of their physiological parameters.

When measurements of physiological parameters are performed using conventional measurement techniques which include procedures like handling, immobilizing or anesthesia , the results must be interpreted with caution as these conventional techniques also have effects on the animals [3]. Radio-telemetry provides a method to obtain accurate and reliable physiological measurements from conscious, freely moving animals [3,4]. It can be used to obtain objective data on acclimatization time and as a tool for defining accurate stabilization periods for laboratory rats following transportation.

By getting more information about transportation-stress, we aim to decrease the variation in research-results and thereby decrease the number of animals needed to get significant results. Secondly, we aim to increase the welfare of laboratory rodents during and after transportation.

Current research

The current research involves both physiological and behavioural measurements in laboratory rats before and after van transportation. Data acquired with bloodpressure-transmitters are blood-pressure, heart rate, respiratory rate and activity. Besides these parameters, home-cage behaviour, bodyweight and faecal [1,6] and plasma-corticosteron were measured. Specific scientific questions that were asked before executing this research were:

- What is the difference in physiology and behaviour in laboratory rats before and after van transportation?
- What is the difference in physiology and behaviour in transported laboratory rats compared to non-transported rats?
- How long does it take an animal to restore these parameters?

Preliminary telemetry and behavioural data will be presented during this presentation

---

References

1. Bamberg, E., Palme, R., and J. G. Meingassner (2001). Excretion of corticosteroid metabolites in urine and faeces of rats. *Laboratory Animals*, **35**, 307-14.
2. Capdevila, S., Giral, M., Ruiz de la Torre, J.L., Russell, R.J., Kramer, K. (2007). Acclimatization of rats after ground transportation to a new animal facility. *Laboratory Animals*, **41,** 255-61.
3. Kramer, K., Acker van, S.A.B.E., Voss, H.P., Grimbergen, J.A., Vijgh van der, W.J.F., Bast, A. (1993). Use of telemetry to record electrocardiogram and heart rate in freely moving mice. *Journal of Pharmacological and Toxicological Methods*, **30-4**, 209-15
4. Kramer, K., Kinter. L.B. (2003). Evaluation and applications of radiotelemetry in small laboratory animals. *Physiological Genomics*, **13-3** 197-205.
5. Obernier, J. A. and baldwin, R. L. (2006). Establishing an appropriate period of acclimatization following transportation of laboratory animals. *ILAR journal*, **47-4**, 364-69.
6. Pihl, L. and J. Hau. (2003) Faecal corticosterone and immunoglobulin A in young adult rats. *Laboratory Animals*, **37**, 166-71.
7. Ruiven van, R, Meijer, G. W., Zutphen van, L. F. M., and Ritskes-Hoitinga, J. (1996) Adaptation period of laboratory animals after transport. S*cand J. Lab. Anim. Sci.*, **23**, 185-90.
8. Ruiven van, R., Meijer, G. W., Wiersma, A., Zutphen van, L. F. M., Baumans, V., Ritskes-Hoitinga, J. (1998) The influence of transportation stress on selected nutritional parameters to establish the necessary minimum period for adaptation in rat feeding studies. *Laboratory Animals*, **32-4**, 446-56.
9. Sánchez, O., Arnau, A., Pareja, M., Poch, E., Ramírez, I., and Soley, M. (2002). Acute stress-induced tissue injury in mice: differences between emotional and social stress. *Cell, Stress & Chaperones*, **7**, 36-46.
10. Swallow, J., Anderson, D., Buckwell, A.C., Harris, T., Hawkins, P., Kirkwood, J., Lomas, M., Meacham, S., Peters, A., Prescott, M., Owen, S., Quest, R., Sutcliffe, R., Thompson, K., (2005).

Guidance on the transport of laboratory animals. *Laboratory Animals*, **39**, 1-39.

11. Tuli, J. S., Smith, J. A., Morton D. B. (1995). Stress measurements in mice after transportation. *Laboratory Animals*, **29-2,** 132-38.

Inline registration of heart rate and body temperature of free swimming eel

Marien Gerritzen, Bert Lambooij, and Hans van de Vis
Animal Sciences Group of Wageningen University and Research Centre, P.O box 65, 8200 AB Lelystad, The Netherlands, marien.gerritzen@wur.nl

Introduction

Production of farmed fish is world wide undergoing a strong growth. Together with this, rearing systems are becoming more intensive and questions concerning animal welfare are raised. Animal welfare aspects of fish are in generally scored based on production parameters (growth, feed consumption) and blood parameters related to neuro-endocrine stress responses. To come to valid statements on fish welfare it is essential to combine production parameters with physiological and behavioural parameters. However, measuring physiological parameters in free swimming fish is at the moment a new field of research and needs development. Intention of this project was to develop a method to measure heartbeat and body temperature in free swimming eels and to correlate these physiological data with activity.

Methods

To monitor on-line heart rate, activity and body temperature continuously for a period of months we implanted a Telemetric device (Telemetronics PhysiolinQ®) in eels. The implant was placed in the abdomen of the eels under full anaesthesia.

Eels of approximattly 40 cm in length and 5 cm in diameter were individually housed in a 100x70 cm tube. To provide surgery eels were placed in water with 0.9% 2-fenoxyethanol to induce full anaesthesia. Anaesthesia was remained during surgery by placing the head of the eel under water with 2-fenoxyethanol. The transmitter (L 2.5 cm x ∅ 1 cm) was placed in the abdome trough a 3 cm inciscion approximatly 3 cm caudale of the liver, wich is approximatly 7 cm caudale of the pectoral fins.

Two leads to measure heart rate where subcutaneous brought to the base of the left and right pectoral fins and secured with a double non-resolvable suture. The abdomen was closed by single monofilament sutures and a thin layer of skin bond. After surgery the eels were placed back into their home tub.

The signal receiver, which also charges the implant, was located under the bottom of the tub.

Recordings

Over a period of 2 months stress was induced by frequently placing other eels in the same tub and by chasing the eels with a net. During these, potentially stressful, interventions heart rate increased together with the activity of the eels. Also body temperature increased for short periods during increased activity.

Measuring and registration of changes in heart rate, body temperature and activity without human intervention is well possible in free swimming eels. Eels showed low activity when they were not stressed and therefore, a bio-rhythm could not be observed.

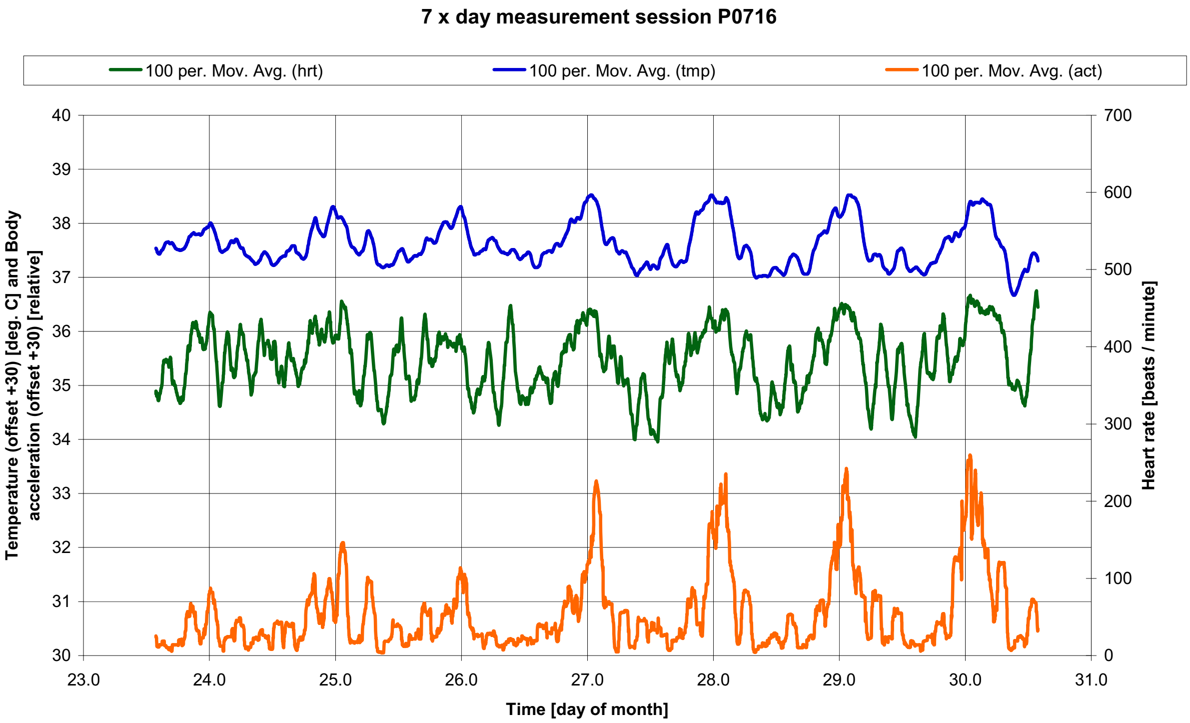


Figure1. *Example off measured body temperature (offset +20), heartbeat and activity (offset+30).*

An integrated system for video and telemetric recording: measuring behavioural and physiological parameters

Elena Moscardo[1], Chiara Rostello[2], Eric Rieux[3], and Reinko Roelofs[4]

[1] Safety Pharmacology Group, Dept. Safety Assessment, and [2]Animal Model Investigator Support Group, Laboratory Animal Science, GlaxoSmithKline R&D Centre, Verona, Italy, elena.2.moscardo@gsk.com and chiara.b.rostello@gsk.com

[3]Eric Rieux, Data Science International, St. Paul, USA, erieux@datasci.com

[4]Reinko Roelofs, Noldus Information Technology, Wageningen, The Netherlands, R.Roelofs@noldus.nl

Physiology gives insight into phenomena that are hard to see, such as sleep changes, seizure electrical activity and memory processes; therefore it is often the combination of physiology and behaviour that makes research most interesting and valuable. The aim of our research activities in the last two years was the setting up and assessment of a new Video-Telemetry system to be applied in pre-clinical research. In particular, our objective was to find an integrated equipment to properly analyze the relationship between electrical brain activity and behaviour in rodents, in order to powerfully assess specific activities of the central nervous system (i.e., seizure, sleep and memory) and their phenotypical-behavioural manifestations. Radio-telemetry is a technology of radio-signals recording from chronically implanted transmitters in freely moving animals widely used to investigate physiological parameters, while still allowing the animal to display natural behaviour and therefore the combination of telemetry and video recording was selected for setting-up our Video-Telemetry system. This integrated system should reduce the number of animals used, improve the data quality by maximizing the amount of information gained from each experimental animal and reduce animal stress and discomfort.

In our laboratory was installed and technically verified for its proper functionality a complete telemetric (by Data Science International) and video (by Noldus Information Technology) system for 8 rats singly housed in PhenoTyper cages, see Figure 1. The selected equipment (hardware and software) allows the simultaneous and continuous recording of physiological parameters (e.g., EEG, EMG, Temperature) and of video images for up to 24-48 hours or longer periods (few days).

The telemetric waveforms recorded with the DQ ART system (DQ ART Gold 4.1, DSI) are then analysed using dedicated software (e.g., Sleep-Sign by Kissei America, Neuroscore by DSI) and the perfectly synchronized video images are shown simultaneously on the screen for helping the researcher during the waveforms analysis. The system also allows using videos for a more detailed analysis of the behaviour, using dedicated software (i.e., The Observer XT and Ethovision XT by Noldus) and the perfectly synchronized telemetric traces are shown simultaneously on the screen for helping the researcher during the behavioural analysis. The analysis of the 24 hours video files with Ethovision XT consists in an initial automatic score of standard behaviours and general activity and then, in the identification of specific episodes based on behaviour. This automatically recorded data can reduce the workload and increase the accuracy of behavioural scoring. A following more detailed manual analysis of the videos is conducted with The Observer XT, that can synchronize multiple behavioural video episodes scored with Ethovision XT with the physiological telemetric data. Male CD™ rats, approximately 7-8 weeks old, were selected for the surgical instrumentation with a telemetric transmitter (type TL10M3-F50-EEE, DSI) implanted intra the peritoneum. The two leads for electroencephalography were placed one on the dura for cortical EEG recording and one deep to the dentate gyrus for the hippocampal EEG recording; the third couple of leads was secured to the neck muscle for the EMG recording [1].

With the common aim to do better science we would share with you our experience on investigating and assessing the potentialities of this integrated DSI and Noldus system to be used for combining and synchronizing the physiological parameters and behavioural patterns.

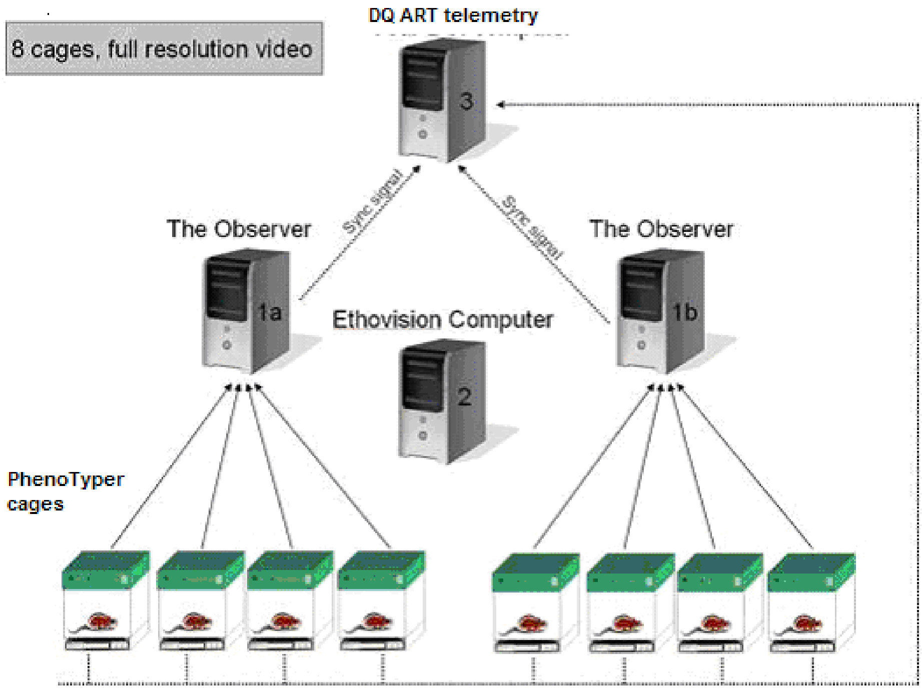


***Figure 1.** The integrated Video-Telemetry system*

- *2x The Observer XT with External data and multiple video modules and multi MPEG-4 encoder*
- *1x EthoVision XT software with multiple body point module*
- *8x PhenoTyper cages for rat with top camera unit*
- *1x Data Quest ART Gold 4.1*
- *2x synchronizations signals C12V analogue to digital converters*

References

1. Williams P.A., White A.M., Ferraro D.J., Clark S., Taley K. and Dudek F.E. (2006). The use of radiotelemetry to evaluate electrographic seizures in rats with kainite-induced epilepsy. *Journal of Neuroscience Methods*, **155** 39-48, 2006.

Concomitant assessment of heart rate and behavior in freely moving mice

Oliver Stiedl, Anton W. Pieneman, Jiun Youn, and René F. Jansen
Center for Neurogenomics and Cognitive Research, Vrije Universiteit Amsterdam, The Netherlands, oliver@cncr.vu.nl

Heart rate (HR) and its adjustment are affected by different physiological conditions such as sleep, physical activity and the emotional state. The instantaneous beat-by-beat fluctuation of heartbeats provides important information about the physiological state of the brain. While the precise origin of beat-by-beat HR variability has not been identified clearly, measures of HR dynamics provide critical clues for the assessment of physiological versus pathological states. By combining autonomic measures in mice during fear learning and memory experiments with advanced non-linear data analysis, we developed a combination of methods that is useful to compare mouse and human HR dynamics for a translational approach [1]. The translation of results is a prerequisite for the development of adequate animal models of affective disorders to improve our understanding of causal mechanisms involved in the comorbidity of emotional and cardiac disorders.

Current telemetry systems are able to monitor ECG, blood pressure and body temperature in mice. However, many commercially available systems provide only averaged HR values. We have mainly focused on HR measurements derived from ECG recordings since (i) HR can be obtained with high precision at high data sampling rates, and (ii) the dynamical adjustment of HR, as opposed to blood pressure, is under direct neuroautonomic control with fast adjustment dynamics in both mice and man. A first important issue is to determine the normal physiological range of HR of mice under baseline stress-free conditions in the home cage. The impact of different recording techniques and experimental conditions on HR is described elsewhere [2]. Long-term HR recordings reveal baseline HR levels in C57BL/6J and C57BL/6N mice predominantly in the range of 550-600 beats per minute [bpm]. With these values in mind it is not surprising that the interpretation of effects of pharmacological interventions with baseline HR values of 720 bpm have led to the misconception that HR in mice is not under parasympathetic control while the sympathetic tone predominates, while our results show exactly the opposite. Experiments in rodents so far have failed to convincingly show the effect of the anxious state on HR dynamics because the experimental conditions under which HR changes have been measured were generally inadequate to assess phasic changes. Any intervention with the experimental animal immediately before the actual anxiety test will confound the subsequent autonomic measure of interest. Thus, behaviorally established anxiety tests such as the dark-light test, elevated plus maze, open field test and novelty exposure are unsuited for measurement of fast HR adjustments because they require handling. Any exposure to a novel environment including a plain change of the home cage serves as unconditioned stressor and results in initial HR values close to 800 bpm representing the maximum physiological limit in this species. The recovery of HR to baseline values (550-600 bpm) is assumed to require 1-2 hr in C57BL/6J mice [3]. While novelty exposure is useful to assess maximum HR values under physiological conditions, it is not expected to reflect HR changes that would occur normally when starting out from baseline values. Therefore it is crucial to change strategies by offering the animal a choice for exploring a novel environment from a home base without interference by the experimentalist. A first attempt has been made in this direction, however, without being able to remove all aversive experimental interference as indicated by highly elevated (> 700 bpm) baseline HR values [4]. Based on this complication and the need for refined experimental conditions, our focus so far has been the investigation of HR responses during expression of fear conditioned to an auditory cue that can be tested in the home cage in combination with genetic or pharmacologic interventions [5-7]. Re-exposure of mice to a tone serving as conditioned stimulus that has been previously paired with aversive stimulus such as foot shock elicits a fear response that results in a profound increase of HR close to maximum physiological limits of ~800 bpm [1,5-7]. Furthermore, startle experiments have been performed under habituated conditions rather than under semi-restrained conditions in conventional startle systems that show that startle responses elicit phasic HR changes that do not suggest a strongly aversive effect [2]. The combination of behavioral and autonomic measurements without intervention by the experimenter is required to further characterize the relation between HR and physical activity in various behavioral aspects in a more thorough way.

A dysfunctional central autonomic system is expected to contribute to elevated risk of cardiac mortality. While the central autonomic pathways have been anatomically characterized in considerable detail, their functional significance is currently still poorly understood because of (i) non-physiologic experimental conditions, (ii) inadequate and insensitive analytical methods, and (iii) lack of spatial and temporal precision of interventions. The appropriate combination of state-of-the-art methods applied in basic research using valid animal models will improve our understanding of the functional significance of central autonomic pathways and their receptors for a translational approach to identify mechanisms underlying human cardiac risk in affective disorders. This translational approach is expected to improve diagnostics and facilitate therapeutic interventions in the clinical setting eventually providing for a stronger linkage between cardiology, neuroscience, psychology and psychiatry (neurocardiology). Aim of the presentation is to provide an overview on current methods and techniques and to give a perspective for future needs in combining behavioral and autonomic measurements.

References

1. Meyer, M., Stiedl, O. (2003). Self-affine fractal variability of human heartbeat interval dynamics in health and disease. *European Journal of Applied Physiology*, **90**, 305-316.
2. Tovote, P., Meyer, M., Pilz, P.K.D., Ronnenberg, A., Ögren, S.O., Spiess, J., Stiedl, O. (2005). Dissociation of temporal dynamics of heart rate and blood pressure responses elicited by conditioned fear but not acoustic startle. *Behavioral Neuroscience*, **119**, 55-65.
3. Stiedl, O., Tovote, P., Ögren, S.O., Meyer, M. (2004). Behavioral and autonomic dynamics during contextual fear conditioning in mice. *Autonomic Neuroscience: Basic and Clinical*, **115**, 15-27.
4. Depino, A.M., Gross, C. (2007). Simultaneous assessment of autonomic function and anxiety-related behavior in BALB/c and C57BL/6 mice. *Behavioural Brain Research*, **177**, 254-260.

5. Tovote, P., Meyer, M., Ronnenberg, A., Ögren, S.O., Spiess, J., Stiedl, O. (2005). Heart rate dynamics and behavioral responses during acute emotional challenge in corticotropin-releasing factor receptor 1-deficient and corticotropin-releasing factor-overexpressing mice. *Neuroscience*, **134**, 1113-1122.
6. Meyer M, Stiedl O (2006) Fractal rigidity by enhanced sympatho-vagal antagonism in heartbeat interval dynamics elicited by central application of corticotropin-releasing factor in mice. *Journal of Mathematical Biology*, **52**, 830-874.
7. Stiedl, O., Misane, M., Koch, M., Pattij, T., Meyer, M., Ögren, S.O. (2007). Activation of the brain 5-HT2C receptors causes hypolocomotion without anxiogenic-like cardiovascular adjustments in mice. *Neuropharmacology*, **52**, 949-957.

Trends in telemetry monitoring: more data and improved animal welfare

Eric Rieux
Data Sciences (DSI), Les Angles, France, erieux@datasci.com

Implantable telemetry has become the gold standard to obtain physiological data from unstressed, free-moving animal models. A wide range of vital parameters such as Heart Rate, Blood Pressure, EEG, ECG, EMG, Temperature, Motor Activity, etc. can be monitored in animals ranging from small mice of 17 grams, to rats and similar size rodents and other larger animal models (dogs, primates, pigs, etc.).

In addition to physiological data, a number of environmental and behavioral data can be simultaneously collected: room conditions such as lighting and temperature, and events such as running wheel, drinking, and feeding activity. External stimuli (light, noise, air jet, etc.) can also be recorded together with the physiological data. And last but not least, valuable video data can be acquired together with the physiological data in order to give a "full picture" of the animals to the investigator.

This presentation will include a short review of telemetry systems, and present three recent technical improvements enhancing animal welfare and social behavior: large animals housed in pair or groups in larger cages with a Telemetry Repeater; pair housed rats in Central Nervous System studies; External telemetry systems for large animal models in Toxicology studies.

Large animals housed in pair or groups in larger cages with a Telemetry Repeater

The Multi-frequency Telemetry Repeater allows studies to be conducted in a group housed environment and/or when there is a need to have telemetry transmitted over longer ranges. The Repeater system allows for greater flexibility in cage setup and housing designs as receiver placement and cage size is less of a concern versus using the standard DSI telemetry receivers.

The Repeater is a device that functions as a signal relay between a DSI Physiotel® transmitter and a Repeater Receiver for large animal and primate telemetry applications. In short, the Repeater receives a signal from an implanted transmitter and retransmits that signal at a different frequency to a Repeater Receiver. The Repeater transmits up to distances of 10 meters.

The repeater is also ideal for toxicology studies. Reduce surgical risk in dogs, primates and other species by using a small animal transmitter in conjunction with the telemetry repeater. When used together, chronic high quality blood pressure and ECG signals can be obtained with a minimally invasive procedure.

Pair housed rats in Central Nervous System studies

The 4ET transmitter was designed primarily for studies involving the Central Nervous System (CNS) in rats and other similarly sized species. It has the ability to monitor up to 4 biopotential channels including any combination of electroencephalogram (EEG), electromyogram (EMG), and electrocardiogram (ECG), along with temperature and locomotor activity. These parameters allow the assessment of sleep architecture, seizure activity, behavior, and other neurological disorders. CNS assessment is particularly challenging in animal models and it is important to maintain a comfortable and controlled environment. The 4ET transmitter allows monitoring of freely-moving animals in their home cage. In addition, it is DSI's first transmitter that allows pair housing to promote more natural, social conditions. Many CNS studies require long-term monitoring, such as those involving aging and cognition. The battery component of the 4ET (telemetry module) can be replaced in-vivo through a minor surgical procedure to prolong the use of the animal. The 4ET transmitter brings improved animal welfare and new research opportunities to the CNS field.

External telemetry systems for large animal models in Toxicology studies

The Jacketed External Telemetry (JET™) Platform is specially designed to work in conjunction with the Ponemah® software for use in toxicology and safety pharmacology laboratories running large animal studies. JET may be used to collect stress-free high fidelity ECG's, respiration, temperature, and activity waveforms from freely moving animals. Up to 36 devices may be used in the same room without interference and no possibility of crosstalk through a specially designed implementation of Bluetooth®. Furthermore, the system is optimized for portability as there is a minimum amount of supporting hardware and it requires no technical expertise to setup. Each device includes an environmentally-friendly removable rechargeable battery to minimize ongoing replacement costs. The final results are hours to days of high quality data obtained quickly and easily and without the time needed for surgery and recovery

References

1. El Amrani Francine, Eric Martel, Pascal Champeroux, Serge Richard, Eric Rieux and Dan Singer (2007). Evaluation of multi-frequency repeaters for record of cardiovascular parameters in group housed dogs. *Journal of Pharmacological and Toxicological Methods*, **56(2)**. September-October 2007, Page e21 *doi:10.1016/j.vascn.2007.02.043*

Improvement of postoperative pain by ropivacaine: a radiotelemetric study of freely-moving rats following calibrated laparotomy

A. Charlet, J.L. Rodeau, and P. Poisbeau

Institut des Neurosciences Cellulaires et Intégratives, Department Nociception and Pain, UMR 7168 Centre national de la Recherche Scientifique - Université Louis Pasteur, Strasbourg, France, alexandre.charlet@neurochem.u-strasbg.fr

Pain is defined by the I.A.S.P (International Association for Study of Pain) as an unpleasant sensory and emotional experience associated with actual or potential tissue damage, or described in terms of such damage. Today, most of the "pain reaction" measured in the animal is based on the analysis of threshold values to produce motor reflexes after noxious or non-noxious stimuli. If useful, these tests require animal handling, produce a variable degree of stress and do not allow to obtain information on spontaneous pain expression [1].

To avoid or limit stress in animal, we have used radiotelemetry to measure some key physiological parameters in freely-moving animals recovering from surgery [2] Transmitters used in the present study (volume: 7cc; weight: 3g) allowed to record changes in the heart rate, respiratory frequency, abdominal temperature and locomotor activity in real time and for the whole duration of the experiments. A surgery is required to implant this type of miniaturized electronic biocaptors in the peritoneal space of the animal and we defined strict conditions of animal surgery time and anaesthesia. To evaluate postoperative pain, we characterized the mechanical sensitivity around the abdominal scar using calibrated von Frey filaments. We first show that, immediately after surgery, the animal exhibited a strong mechanical allodynia, *i.e.* animals show a pain symptom although mechanical stimulation is non noxious. This mechanical allodynia in the periphery of the abdominal scar persists significantly for up to 10 days. Interestingly, pain hypersensitivity is timely associated with a significant locomotor impairment, hyperthermia and tachycardia.

To confirm which of these symptoms are correlated with pain expression, we characterized the effect of ropivacain, a long-lasting local anesthetic [3] on the postoperative recovery. If ropivacain is administered subcutaneously around the scar either before the beginning of surgery or just before the awakening of the animal [4], no more signs of mechanical allodynia could be detected during the following days after surgery. Locomotor activity and heart rate values, measured in ropivacaine-treated animals, were similar to those of animals having fully recovered (*i.e.* more than 15 days after implantation). Interestingly, only hyperthermia persisted 7-10 days and was unaffected by the ropivacaine treatment, letting us thinking than that is induced by scaring process, and do not reflect pain. This was not the case if the single infiltration of ropivacaine was performed 3 days after the surgery. Indeed, we observed a transient and short-lasting analgesic effect but we did not see any changes in the locomotor impairment, tachycardia or hyperthermia.

In summary, we have been able to monitor simultaneously the pain hypersensitivity resulting from abdominal surgery and some physiological/autonomic parameters. This study was conducted in freely-moving and non-handled animals which limit possible bias due to stress. We have characterized the time-course of mechanical allodynia and the impairment of locomotor and heart rate values during the recovery period of the surgery. Strongly supporting than control of side-effects is crucial for successful surgery and fast-track rehabilitation [5] we show than modifying slightly the surgery protocol by using subcutaneous ropivacaine injections in a short time window (*i.e.* before or just before awakening of the animal) appears to prevent the immediate and long-lasting allodynia as well as locomotor impairment and tachycardia. The mechanism involved in this phenomenon is currently under investigation.

References

1. Le Bars, D., Gozariu, M., Cadden, S.W. (2001). Animal models of nociception. *Pharmacol. Rev.*, **53**, 597-652.
2. Kramer, K. & Kinter, L.B. (2003). Evaluation and applications of radiotelemetry in small laboratory animals. *Physiol. Genomics*, **13**, 197-205.
3. Hansen, T.G. (2004). Ropivacaine: a pharmacological review. *Expert Rev. Neurother.*, **4**, 781-791.
4. Moiniche, S., Kehlet, H., Dahl, J.B.(2002). A Qualitative and Quantitative Systematic Review of Preemptive Analgesia for Postoperative Pain Relief. *Anesthesiology*, **96**, 725–41
5. Kehlet, H. & Dahl, J.B. (2003). Anaesthesia, surgery, and challenges in postoperative recovery. *Lancet*, **362**, 1921-1928.

Measuring consumer behavior Special Interest Group

Hans Theuws
Noldus Information Technology bv, Wageningen, The Netherlands, H.Theuws@noldus.nl

Introduction

Observation of consumers in specific situations is becoming increasingly important in predicting the success of new products. Observing consumer behavior and the reactions triggered by such behavior provides new possibilities and opportunities for improving consumer products. Researchers are more and more aware of the many advantages new technology has to offer for measuring consumer behavior. A growing number of techniques are available to assist them during observation and measurement of purchasing, selection and consumption behavior. Due to advances in digital video, sensor technology and computer speed, complex measurements of behavior and physiology are now possible. Integration of these techniques allows multimodal measurements. With the growing number of techniques, the challenge for the researcher to choose the right solution becomes bigger. Questions to be answered in order to select the right combination of solutions include:

How do I create the optimal test environment? In what context do I measure consumer behavior? In the recently constructed 'Restaurant of the Future' in Wageningen (The Netherlands), observation of everyday behavior can be combined with physiological and sensory measurements. Other research requires observation of consumers in other situations, for examples a hospital, supermarket, school canteen, etc.

Which tools do I use for observation: video, physiological measurements, facial expressions analysis, head movement, eye-tracking – or a combination of these modalities?

Measuring consumer behavior can result in an overload of acquired data. How do I select and analyze these data in an effective way?

This SIG meeting intends to bring together users and developers of measurement tools for consumer behavior, to discuss the state of the art, advancements, experiences, expectations and bottlenecks. As such it will provide a platform for exchanging information about the opportunities, challenges and needs in the fast developing area of consumer behavior research. It will be organized as a series of presentations, followed by group discussion.

Audience

This SIG aims to bring together researchers involved in consumer behavior research, as well as developers of technology and tools for measuring consumer behavior.

Special Interest Group Contents

Innovative Consumer Studies at the Restaurant of the Future

H.E. Schepers, R. de Wijk, J. Mojet, and A.C. Koster

The influence of lighting in the build environment: a study to analyse human behaviour and perception as measured by mood and observation

K. Quartier and K. Van Cleempoel

ConsuNautTM: Consumer analysis in place and time

P. Vasara, H. Lehtinen, and P. Kotro

Advances in face and gesture analysis

H. van Kuilenburg, M.J. den Uyl, M.L. Israël, and P. Ivan

Inside Consumer Experience: Mobile Laboratory to Study Consumer Behavior in the Field

Leanne Loijens

Innovative consumer studies at the Restaurant of the Future

H.E. Schepers, R. de Wijk, J. Mojet, and A.C. Koster
Centre for Innovative Consumer Studies, Wageningen University and Research Centre, Wageningen, The Netherlands, Hans.Schepers@wur.nl

Abstract

Consumer research on eating & drinking behaviour suffers from several drawbacks, such as its reliance on verbal methods (questionnaires, focus groups, interviews), unnatural test environments and professed rather than actual preferences and likings. Attitudes are commonly thought of as personal state variables, even when it has been widely recognized that the correlation of intentions / attitudes with behaviour is poor (see Köster, 2003). Alternative approaches rely on direct non-invasive observation of actual behaviour, and measurements of behavioural variables such as walking, food choices, eating, and talking) in natural environments over longer periods of time.

Assessing spontaneous behaviour

The Restaurant of the Future in Wageningen, The Netherlands, houses, apart from standard sensory and physiological laboratories also a restaurant for daily lunches for employees of Wageningen University & Research Centre. In this facility, eating & drinking behaviour, as well as the selection process of the daily lunch is studied, using a variety of sensors. Lunch guests are only asked for demographic data when they register as customer of the Restaurant and where they agree that their data may be used for scientific purposes. During the years to come, the lunch users will not be addressed by the research team anymore, as their *spontaneous* behaviour is to be investigated, and it has become clear that attention is an all too powerful factor influencing food behaviour. Wansink & Sobal (2007) documented how various situational factors, such as odours, lighting, and seating arrangements influence food behaviour sub-consciously. We will further investigate the causal pathways behind these phenomena, as well as their time scales of change. Ultimately, we wish to integrate the findings into a system dynamics style model that explains and predicts longer term behavioural change as a function of the foods offered, situational factors, such as the physical and social context and the person's psychological and physiological states.

In order to monitor the purchase behaviour of the consumers in the lunch restaurant, we keep records of all cash registers data in Point of Sales (POS) database. Even when several products such as various types of soup are priced identically, we keep unique articles codes for them, in order to distinguish vegetarian, thick or clear, spicy or creamy soups etc. Although standard practice in a retail environment, this is not the case in a food service context. However, food service POS data can potentially lead to much deeper consumer insights in food preferences, as the customers who buys the food is typically also the person who consumes the food in a well-defined environment. This is in contrast to retail purchasing records, which reflect household needs for several days.

Although technically not innovative, we will be able to follow consumers' reactions to various situational and product-offering manipulations at a daily basis, yielding a rich data set to link all other measurement records to.

Next to the POS data, the Restaurant of the Future is equipped with 26 dome camera's in the ceilings that registers visible behaviours related to entering, searching for and picking all food items, selecting a table and finally eating lunch. Also disposals of uneaten foods and packaging can be obtained, assessed and monitored. Video recordings are coded for specific behaviours using Noldus Observer. Correlating overall and partial residence times as well as group dynamics (eating alone or in a group) with POS data can reveal how time pressures, social context co-vary with purchase behaviour. These outcomes are collected over longer periods of time and can be correlated with season, weather patterns etc. Clearly, such conclusions yield valid insights in behaviour, that would not be obtained by asking people "what would you eat when you are not in a group during winter?".

Semi-invasive measurements

The RoF can add monitoring physiological measures of selected consumers, using heart rate, skin conductance, weight, blood pressure registrations etc. If these measurements interferes with the spontaneous behaviour of other customers, these half-invasive studies will be performed either at other times or at other locations in the Restaurant. In such situations more close measurements on physiology can be conducted on persons prior or after selected and having eaten certain products, in order to study e.g. satiation and satiety.

Similarly various devices can be used to measure important variable such as attentional focus, such as using eye tracker software and hardware, in order to see what consumers are looking at, which signs they might notice or read, which meal presentations are seductive etc etc. In addition, facial expressions, head directions can be interpreted from video recordings, using Facereader software. Ideally, one would uncover stronger correlations with such body language and food behaviour, relative to possible patterns one could try to establish comparing food related behaviour to verbal language answers and opinions.

Finally, integration of long term purchase patterns, food perception, expectations and memory under various circumstances, as well as assessment of eating speed and bite size (as assessed through coding video recordings) will be vital to support new product development and product introduction in the out of home markets.

References

1. Koster, E.P. (2003) The psychology of food choice: some often encountered fallacies. *Food Quality and Preference*, **14**, 359–373.
2. Wansink, B & J. Sobal 2007: Mindless eating: The 200 daily Food decisions we overlook: *Environment and Behavior*, **39**, 106-123.

The influence of lighting in the build environment: a study to analyse human behaviour and perception as measured by mood and observation

K. Quartier and K. Van Cleempoel

Department of architecture, art and design, PHL University College, Hasselt, Belgium, katelijn.quartier@phl.be

Abstract

My research involves understanding human movement through lighting. For ages people have been attracted by light, either by fire, candles or more recently lamp bulbs. Out of this point of view the question arose whether humans are influenceable in the way they move in the build environment, retail environments in this case, guided through lighting. Hence, the first focus of the study lies on how humans behave (move) under certain lighting conditions: way-finding as well as their walking speed are included. The second focus of this study lies more in how humans perceive that space under certain lighting conditions, on a more emotional based approach. The experiments are currently and in the next few months held in a controlled environment, which is a simulation of a supermarket. In a next stage the results will be validated in real settings. Together with the theoretical background, already developed, the results of these experiments will give a clear view on the how humans move through space and how they perceive space, by the manipulation of lighting.

Introduction

The goal of this research is the development of guidelines for designing lighting in the build environment. Two statements lie on the bases of this topic. Firstly, lighting can be designed to guide people's movement through space. Secondly, lighting can be designed in such a way it has a positive influence on the appreciation of a space, via the perception of that space. To back these statements up, a theoretical framework has been developed. This framework is based on research in commercial spaces and workspaces since lighting has a huge impact on humans in both cases. The choice to undertake this research in commercial spaces is bilateral: a thorough research indicated that commercial spaces are ideal environments to experiment in with lighting, humans, their behaviour and emotions; furthermore, since most research has been done in office settings and there is a gap in research for retail settings, it is only logic to develop hypotheses -for retail settings- based on what is already been studied –in office settings.

Regarding retail environments, the following has been studied. Research showed that shop environments create a 'retail experience' that strongly influences consumers' purchase behaviour [1-2]. Moreover, keeping shoppers longer in stores is likely to result in increased browsing behaviour, which in turn is likely to cause increased impulse purchasing [3]. And even two third of purchase decisions are made in store [4]. Consequently, shop interiors are extremely important and through their interior variables have an individual effect on the consumer. Kotler [5] therefore, coined the term '*atmospherics*' to describe "*the effort to design buying environments to produce specific emotional effects in the buyer that enhance his (/her) purchase probability*". Many design elements in combination or separately exert this influence on behaviour in general. In this paper we want to limit our research to the factors 'light', the perception of space and movement trough that space. How is light influencing behaviour (movement and perception), and what kind of behaviour precisely are we talking about? Bitner [6] analysed how consumers respond to a retail environment and noted that consumers can react to a retail store in a cognitive, emotional and biological way. Our research aims at two of these aspects: on one hand, the emotional factor, this is the effect of lighting conditions through our perceptual system. The physiological factor, on the other hand, is also included as lighting also has a purely biological influence on people. Discussed here are the short term biological effects of lighting on people – the way people behave under certain light conditions, which route they take and with which speed- not the long term biological effects - which imply the shift of the body's circadian rhythms. Both the emotional and physiological factors are hard to separate and they will influence one another.

There are three parts in this study and presently the second part is being undertaken. The first part included a survey of the literature, replenished with in depth interviews with selected experts. As a result of this part, certain hypotheses were developed (point 3). The second part is the experiment in a controlled environment –supermarket- and subsequently, the third part will be the development of guidelines for lighting design in (commercial) spaces, based on the results of experiments in real settings. The linkage between research and design will be taken to the highest possible level in this last phase. The experiments in the lab environment are firm intern valid, so the results of the real retail environment will make it also external valid. Therefore both experiments enhance each other and they will provide results that form a translation of how research can be implemented and translated in the design process. As a designer, extended with research skills (due to the PhD I'm currently undertaking), I put myself in the position to take count of both the users' point of view (costumers journey) and designers' point of view. This results in a balance between the rational (designer) and emotional (user) values of the design process.

Short overview of the theoretical framework

When reviewing the literature regarding the influence of lighting within retail environments, several aspects have been studied. A first aspect is concentrated on consumer behaviour on a very basic level: Taylor and Socov's research [7] showed that light influences the route consumers take through the store. This study indicates that people are drawn to light and that they, therefore, will choose the more illuminated path when passing an obstacle.

A second aspect handles on a more product-based level, which in turn also influences peoples' behaviour. In this context LaGuisa and Perney [8] stated that light could draw attention to products. Areni and Kim [9] added to this knowledge by doing experiments with wine bottles in a store. Under 'bright lighting' conditions bottles were more often examined and touched than under 'dim lighting' conditions. Magnum [10] took this one step further and showed that lighting influences the attractiveness of products in a store. Along the same line, Summers and Hebert [11] showed that more belts were touched and picked up with the addition of display lighting.

Subsequently, they constituted consumers spent significantly more time at the display with the additional accent lighting.

A third aspect regarding light in retail environments comes from an environmental psychology based emotional model. Mehrabian and Russell [12] developed this model to study environmental stimuli within retail environments and their influence on the consumer. They proved that emotions, evoked by shop environments, are related to consumer behaviour, and one step further, to buying behaviour [13-14]. But each of them only make assumptions about the lighting specifically, and remain very vague about its possible influence on consumer behaviour. What we do know is that lighting has an influence on the perceived price level of a store (Luomala, 2003) [15]. What most of us already assumed is scientifically proven: bright lighting conditions, in combination with orange coloured walls induce the feeling of low price perceptions. Soft lighting conditions on the contrary appear to increase the height of the price perception.

The biological influence in retail environments has not received any attention yet. However, the biological influence of lighting in working environments received a great deal of attention. So far, a precise analysis of what exactly light does to the human body remains incomplete. What is pronounced is that the influence of lighting is not unambiguous. Studies in psychology have found that individual personality traits can influence reactions to environments [16]. Furthermore, Knez and Kers [17] found that age and gender interacted with the illuminance and the colour temperature of the lighting, causing different kinds of mood shifts, in working environments. Kuller [18] showed differences in reactions to lighting, based on cultural backgrounds. Also, the spectrum of the lighting plays an important role as well. Studies about this give controversial results [19-23]. These aspects are taken into account in the development op this study.

Hypothesis development

The hypotheses developed are based on the theoretical framework and they are currently tested in a controlled lab environment.

Hypotheses in general spaces are:

Lighting has an influence on the mood and hence the behaviour of humans in (retail) environments.

Lighting has an influence on how spaces are perceived; even in a matter it can arouse positive or negative feelings in the perception of that space.

Lighting has an influence on way-finding as well as ones walking speed in a (retail) environment.

Hypotheses specifically retail orientated are:

Positive effect, induced by lighting in retail environments, will encourage humans to stay longer in that retail environment.

Positive affect, induced by lighting in a retail environment, will increase the sales numbers of that retail environment.

Note here that the persistency of lighting and its effects might be relative to all the other factors that can change behaviour, perception and movement. Lighting is considered as just one of the many interacting factors that determine the outcome. This is taken into account.

To situate this particular research, regarding the influence of lighting on mood and behaviour, it is necessary to mention that it is a part of a bigger research project –that will start in October- that accents more the analysis of lighting in the build environment through cognitive mapping. The environmental psychology based analyses will be translated in appreciation models and colour perception indexes. Colour rendering, colour discrimination and colour harmony under certain lighting conditions are several topics that will be looked into.

Finally, without lighting, there is no space, there is no build environment and even humans are kept in the dark. Let this be a statement to show the importance and the influence of light and lighting design on human behaviour and perception.

References

1. Moye, L.N., & Kincade, D.H. (2002). Influence of Usage Situations and Consumer Shopping Orientations on the Importance of the Retail Store Environment. *International Review of Retail, Distribution and Consumer Research*, **12(1)**, 59-79.
2. Sherman, E., Mathur, A., & Smith, R.B. (1997). Store Environment and Consumer Purchase Behavior: Mediating Role of Consumer Emotions. *Psychology & Marketing*, **14(4)**, 361-378.
3. Beatty, S., & Ferrell, M.E. (1998). Impulse Buying: Modelling Its Precursors. *Journal of Retailing*, **74(2)**, 169-19.
4. POPAI Europe, (1998). *The POPAI Europe Consumer Buying Habits Study*. Point-of-Purchase Advertising Institute. Co-ordination by Retail Marketing In-Store Services Limited, Watford, Herts : POPAI Europe.
5. Kotler, P. (1973). Atmosphere as a Marketing Tool. *Journal of Retailing*, **49**, 48-64.
6. Bitner, M.J. (1992). Servicescapes: The Impact of Psychical Surroundings on Customers and Employees. *Journal of Marketing*, **56(2)**, 57-70.
7. Taylor, L.H., & Sucov, E.W. (1974). The Movement of People toward Lights. *Journal of the Illuminating Engineering Society*, **3**, 237-241.
8. LaGiusa, F.F., & Perney, L.R. (1974). Further Studies on the Effects of Brightness Variations on Attention Span in a Learning Environment. *Journal of the Illuminating Engineering Society*, **3**, 249-252.
9. Areni, C.S., & Kim, D. (1994). The Influence of In-store Lighting on Consumers' Examination of Merchandise in a Wine Store. *International Journal of Research in Marketing*, **11**, 117-125.
10. Mangum, S.R. (1998). Effective Constrained Illumination of Three-Dimensional, Lightsensitive Objects. *Journal of the Illuminating Engineering Society*, **27**, 115-31.
11. Summers, T.A., & Hebert, P.R. (2001). Shedding Some Light on Store Atmospherics: Influence of Illumination on Consumer Behaviour. *Journal of Business Research*, **54(2)**, 145-150.
12. Mehrabian, A., & Russell, J.A. (1974). *An Approach to Environmental Psychology*. Cambridge MA: MIT Press.
13. Moye, L.N., & Kincade, D.H. (2002). Influence of Usage Situations and Consumer Shopping Orientations on the Importance of the Retail Store Environment. *International Review of Retail, Distribution and Consumer Research*, **12(1)**, 59-79.
14. Sherman, E., Mathur, A., & Smith, R.B. (1997). Store Environment and Consumer Purchase Behavior: Mediating Role of Consumer Emotions. *Psychology & Marketing*, **14(4)**, 361-378.
15. Luomala, H.T. (2003). Understanding how retail environments are perceived: a conceptualization and a pilot study. *International Review of Retail, Distribution and Consumer Research*, **13(3)**, 279-300.
16. Russell, J.A., & Snodgrass, J. (1991). *Emotion and the Environment*. In: Handbook of Environmental Psychology, Stokols D., & Altman I., (Eds.), Malabar, FL: Krieger Publishing Company, 245-280.
17. Knez, I., & Kers, C. (2000). Effects of Indoor Lighting, Gender and Age on Mood and Cognitive Performance. *Environment and Behavior*, **32(6)**, 817-831.

18. Kuller, R., Ballal, S., Laike, T., Mikellides, B., Tonello, G. (2006). The Impact of Light and Colour on Psychological Mood: a Cross-Cultural Study of Indoor Work Environments. *Ergonomics*, **49(14)**, 1495-1507.
19. Boray, P.F., Gifford R., & Rosenblood, L. (1989). Effects of Warm White, Cool White and Full-spectrum Fluorescent Lighting on Simple Cognitive Performance, Mood and Ratings of Others. *Journal of Environmental Psychology*, **9**, 297-308.
20. Baron, R.A., Rea, M.S., & Daniels, S.G. (1992). Effects of Indoor Lighting (Illuminance and Spectral Distribution) on the Performance of Cognitive Task and Interpersonal Behaviors: the Potential Mediating Role of Positive Effect. *Motivation and emotion*, **1**, 1-33.
21. Küller, R., & Wetterberg, L. (1993). Melatonin, Cortisol, EEG, ECG and Subjective Comfort in Healthy Humans: Impact of two Fluorescent Lamp Types at two Light Intensities. *Lighting Research and Technology*, **25(2)**, 71–81.
22. Rusak, B., Eskes, G.A., & Shaw, S.R. (1996). *Lighting and Human Health*. A Review of the Literature. Ottawa, Ontario: Canada Mortgage and Housing Corporation.
23. Veitch, J., & McColl, S. (2001). A Critical Examination of Perceptual and Cognitive Effects Attributed to Full-spectrum Fluorescent Lighting. *Ergonomics*, **44**, 255–279.

ConsuNaut™: Consumer analysis in place and time

P. Vasara[1], H. Lehtinen[1], and P. Kotro[2]
[1]*Pöyry Forest Industry Consulting Oy, Helsinki, Finland, petri.vasara@poyry.com*
[2]*Valkeus Interactive Oy, Helsinki, Finland*

An established component of market research is the assessment of people's use of time: how many minutes per day are spent in different activities. New technologies such as sensors and GPS enable collecting data on locations. We present a method which combines the time and location dimensions in analysing the behaviour of a group of people. In addition to time and place, the method allows observation of moment-related moods and attitudes of people. Furthermore, the impacts of external factors can be included in the methodology analysis.

The method is built around the ConsuNaut™ tool which is a program consisting of a MySQL database and a graphic user interface in Flash. The interface presents subject groups as one hundred circle symbols, each one representing one percent of the group. We can cluster subject group members into various segments which are indicated with different colours. The hundred circles are placed on the ConsuNaut grid which has five different locations, four in the corners and one between the others acting as a transition area. The locations can be either physical places or mental states of the subject group members. One grid view shows the situation of an hour, and the tool can present a series of these time slices of selected hours, e.g. each hour of one week or one month. We can add visual symbols on the circles to describe selected characteristics of the subject group members, these can be e.g. state of each percentage which can have options such as sleeping, eager to learn, and awake but not interested. The colour of cells of the grid can be changed to describe the current circumstances in different locations, e.g. if in one location the competition in providing services for the consumers is fiercer than on the other locations, the appropriate cells may have different colour.

Background data for ConsuNaut™ tool can be collected either from statistics, questionnaires, or more innovatively through mobile phone applications, where participants of the research submit their status report regularly through their mobile terminals. A status report contains a couple of indicators depicting their state of mind, mood or some other characteristics relevant to the study being undertaken. Their location can also be registered in the database. A GPS-enabled phone can store this information; with a phone lacking this ability, coding can be used (e.g. store, home, transit). One example of data editing is the case where we can combine two complementary data sets by building an optimisation model in a spreadsheet computation to match the data in the most accurate manner. Another example is that we use Theme (by Noldus) [1] or SOM (Self-Organising Map, a neural network) [2] to find patterns in human behaviour from the data of a group, and based on these regularities we can convert the data to ConsuNaut™ format.

We implement the analysis phase after finishing the dataset compilation. The motions of groups have previously been animated for instance by simulating behaviour of birds by using distributed behavioural model by Reynolds [3] and by simulating urban crowds based on individual and collective geospatial intelligence by Torrens [4]. The ConsuNaut™ tool simulates group motions simultaneously with mental states of a group, based on the data analysis. The tool enables watching all selected time slices in a row as a slide show, or picking one time slice at the time and taking another time slice next to it for a comparison. The main idea of this kind of analysis is to view large amounts of data at a glance and perceive behaviour patterns visually. This makes it possible to use the inbuilt human perception abilities to convey understanding of a very complex issue: the behaviour in space and time of groups of consumers. The objectives of the analysis depend totally on the definition and the scope of the study. The question to be solved can be e.g. understanding what kind of people are in the transition space at different times of the day or week. Another example is to see where and when a company should market its products, where potential customers are and when they are in the right mood.

The analysis can also contain an assessment of external factors or possible trend impacts on the behaviour of people. This type of assessment naturally requires extra calculations with the background data. We create response functions for each subject group category and use them to produce migration coordinates for the population studied – be it coordinates on the 5-area map of ConsuNaut™ or e.g. coordinates of mood. New locations and other characteristics for each time slice are the output, when input is a selected external change. We can compare a certain time slice before and after the external change to see the impacts or we can view the whole time period in a new situation to see if the patterns have changed. In some cases it is also possible to use the above-mentioned SOM to analyse the behaviour transitions.

References

1. Magnusson, M. S. (2004). Repeated Patterns in Behavior and Other Biological Phenomena. In D. Kimbrough Oller and Ulrike Griebel, editors, *Evolution of Communication Systems: A Comparative Approach*, 111-128.
2. Kohonen, T (2001). Self-Organizing Maps Self - Organizing Maps New, extended edition in 2001. Springer Series in Information Sciences **30**. Third Extended Edition, 501 pages.
3. C. W. Reynolds (1987). Flocks, Herds, and Schools: A Distributed Behavioral Model. Symbolics Graphics Division in *Computer Graphics*, **21(4)**, 25-34.
4. ACM SIGGRAPH '87 Conference Proceedings, Anaheim, California, July 1987.
5. Torrens, P.M. (2007). Behavioral intelligence for geospatial agents in urban environments. IEEE Intelligent Agent Technology (IAT 2007). Los Alamitos, CA, IEEE, 63-66.

Advances in face and gesture analysis

H. van Kuilenburg, M.J. den Uyl, M.L. Israël, and P. Ivan
VicarVision, Amsterdam, The Netherlands, info@vicarvision.nl

Recent technological developments are opening new possibilities for consumer behaviour analysis and consumer interaction. Faces are one of the primary channels for humans to transmit emotional signals, such as the level of appreciation or interest. These are most often uncontrolled spontaneous signals which are more honest and instantaneous and thus more useful for consumer behaviour research than self-reported values.

This article describes a number of new developments in face and gesture analysis which allow a more accurate, more extensive and more reliable analysis of faces, plus the ability to determine the point of interest and basic head gestures.

Facial expression analysis

In 2007 VicarVision and Noldus Information Technology launched FaceReader™, a system for fully automatic facial expression analysis. Although the FaceReader system opens many new possibilities for behavioural research, the scope of settings under which FaceReader can be used has its limitations. Future releases will include technological improvements that will further increase the range of use of the system.

What makes automatic facial expression analysis so difficult? The technical answer to this question is: the high dimensionality of the underlying mathematical problem space. Clearly, an image or video of a face is not suitable for facial analysis in its raw form (an array of pixels), so steps must be taken to reduce the raw data into a graspable set of features describing a face before it can be accurately analyzed. Sources of variation that need to be accounted for in order to reduce the dimensionality include: location of the face in the image, size and orientation of the face, lighting of the face, personal variations in for example gender / age / ethnicity and finally the facial expressions we are interested in.

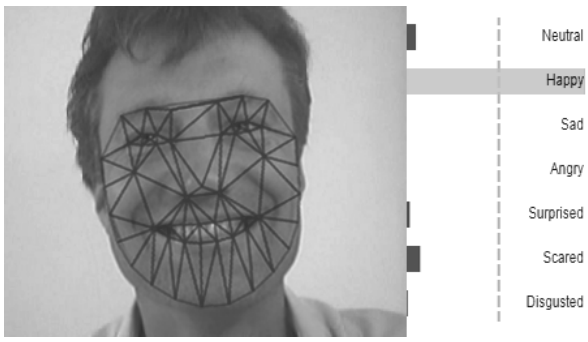
Face detection is the first step in obtaining the important information from a face image or video frame. Several well known methods exist for this task, but each method has some tradeoffs in terms of speed, number of detections and/or framing accuracy. The latest face detection system that will be used in the next FaceReader release uses a unique combination of two face detection algorithms to find faces under a very wide range of variations while still creating a accurate face framing (and in a reasonable amount of time). The popular Viola-Jones algorithm [5] is used to roughly detect the presence of a face after which a deformable template method [4] creates a more accurate framing containing information about the likely in-plane rotation of a face.

The next phase in the FaceReader system is an accurate (photorealistic) modelling of the face using an algorithmic approach based on the Active Appearance method described by Cootes and Taylor [1]. A trained appearance model has limits on the amount of variation that it is able to model, which manifests itself for example as a lower modelling accuracy (or failure) for people of certain ethnicities and age and difficulties modelling faces under certain lighting and orientation. Currently the FaceReader system works best for Caucasian middle-aged people with little facial hair and a frontal camera position and lighting. Dedicated models are being developed for people from Asian origin and for

Figure 1. *Mean faces of FaceReader appearance models*

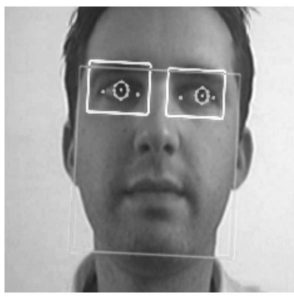
children, which should greatly increase the worldwide usability of the FaceReader system.

In addition to creating more of these dedicated models, a new architecture is being considered which should expand the range of acceptable faces even further. If instead of having a few models, which the user can choose from, we have a (large) number of partly overlapping models, we can create an automatic selection mechanism that would choose the most appropriate model for the current person/settings. The training examples used to create these models could be automatically selected by a mechanism designed to include a certain amount of variation in each model which is known to work well. Increased processing speed and parallel architectures in modern PCs are making it a feasible solution to run multiple models parallel in real-time.

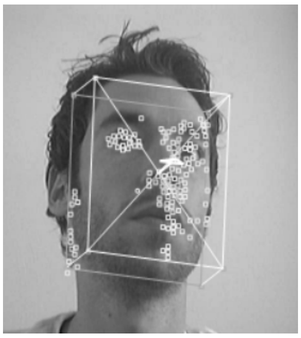


Gaze tracking

In addition to classifying emotional states, determining the object or source of the affect is often of equal importance for consumer behaviour research. In laboratory settings the number of stimuli can be limited, but 'in the wild' it is nearly impossible to determine the causes of all emotional expressions. For example consider a supermarket where a FaceReader system registers a "pleasantly surprised" emotion. Without knowing exactly which product or advertisement the person was looking at when the emotion was registered, this information is of little use.



VicarVision is developing an eye-tracking tool designed to be used with low cost off-the-shelf equipment for viewers that may move around freely over some distance. In contrast to special purpose eye-tracking systems, our system uses standard webcams without infrared capabilities and infrared lighting. Infrared lighting has clear advantages for eye-tracking as the pupil of the eye shows distinct and precise corneal reflections when lit with infrared light. However, standard webcams are easily available everywhere and do not require special lighting that may not work well over larger viewer distances. In a study we showed how the Active Appearance Model is a suitable candidate for creating accurate eye models containing both texture information and positions of key landmark points within the eyes such as the pupil centre and eye corners [2]. This information combined with head position and orientation information allows us to determine the gaze direction with reasonable accuracy.



Robust head orientation tracking

In uncontrolled settings, people tend to look around a lot and make fast head movements which cannot be handled by either the face detection or the face modelling algorithm. The result of this is that the face can only be tracked and analyzed a fraction of the time and all other behaviour of the subject is lost.

To overcome these problems, a module based on the "KLT point tracking algorithm" described by Shi and Tomasi [3], combined with a cylindrical head model is being developed. The point tracking algorithm finds features, within the bounding box of the face, which can be tracked both in the x- and y-direction, such as corner points and checkered textures. In an initial frame these key points are projected on one half of a cylinder (which forms a rough approximation of the shape of a face) with an initial orientation. In following frames the transformation of the cylinder that best fits the translation of the feature points in this frame is estimated. From this change in orientation of the cylinder (and thus change in orientation of the face) can be derived. This technique enables reliable orientation estimation for up to 90° away from the camera and a continuous head tracking in all orientations. Also, head gestures like nodding "yes" or shaking "no" can be recognized.

In addition to more robust face tracking, the face orientation information is also crucial for the gaze estimation module discussed before. In a project with a leading producer of skin-care products, this method also proved to be most valuable for analyzing consumers using products in front of a mirror while not being restricted in head orientation.

Conclusions

Several algorithmic improvements in the FaceReader and the development of new tools will lead to an increased robustness and usability of face analysis technologies . On the long term, significant advances are made to being able to analyze faces in completely uncontrolled settings and extract more information than the emotional expression only. All this is meant to open many new possibilities for consumer behaviour research.

References

1. T. Cootes and C. Taylor. Statistical models of appearance for computer vision. Technical report, University of Manchester, Wolfson Image Analysis Unit, Imaging Science and Biomedical Engineering, 2000.
2. P. Ivan. Active appearance models for gaze estimation. Master thesis, BWI, Free University Amsterdam, The Netherlands, 2007.
3. J. Shi and C. Tomasi. Good features to track. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (CVPR'94), pages 593-600, IEEE Computer Society, Seattle, Washington, June 1994.
4. K.K. Sung and T. Poggio. Example-based learning for view-based human face detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **20(1)**, 39–51, 1998.
5. P. Viola and M. Jones. Rapid object detection using a boosted cascade of simple features. *In proceedings CVPR*, 2001.

Inside Consumer Experience: Studying human food choice in real-life contexts

Leanne W.S. Loijens

Noldus Information Technology BV, Wageningen, The Netherlands, l.loijens@noldus.nl

Introduction

The food and drink industry annually spends large amounts of money to introduce new food products to the market. In spite of all this money, almost 90% of the new products fail [1]. In the laboratory, test panels said they like the products, but in real-life contexts the products fell short. The growing insight that situational influences may be more prominent in the determination of food choice than general overall preferences [2-4], has been the incentive to start the Inside Consumer Experience (ICE) project. The aim of the project is to develop novel instruments and mobile services for the objective measurement of food selection and consumption in real-life contexts.

Food choice

Currently, almost 90% of the new food products that are launched, disappear from the market within one year [1]. One of the main reasons for this high failure rate is the underestimation of the role that situational factors play in food choice [2-4]. Consumer tests are often carried out without taking the particular consumption situations and the related experiences into account. Testing of consumer preferences is primarily done in food laboratories that offer a highly controlled testing environment. The results of these tests cannot be generalized easily to more natural environments.

The gap between laboratory and real-life situations is currently bridged by studies that use typical marketing research techniques such as interviews and questionnaires. The major drawback of these techniques is that they address the consumer's conscious food experience by measuring its verbal expressions. Unconscious processes, nonverbal reactions and important context factors which strongly influence consumer attitudes and behavior cannot be considered.

Figure 1. *Cameras allow you to study eating behavior under natural conditions.*

Restaurant of the Future

Food scientists agree that a new approach to research is needed and that it is necessary to study unconscious consumer behavior. This has led to the design and construction of a highly advanced facility for research on food-related behavior in an indoor setting: the Restaurant of the Future in Wageningen (The Netherlands). The restaurant offers a flexible eating environment where food choice and eating behavior can be observed under natural conditions with the help of state-of-the-art observation and sensor technology (for more information, see www.restaurantofthefuture.nl).

Obviously, not all eating environments can be created or simulated in the Restaurant of the Future. Examples are settings such as daycare centers, elderly homes, or events for teenagers (pop festivals, dance parties, etc.).

Figure 2*. In the control room, the researchers can watch views from several cameras on a big screen.*

Mobile lab

The ICE project takes the concept of situational research a major step further by designing a mobile observation laboratory to study food selection and consumption in real-life contexts, and a suite of corresponding research services. The services which will be offered, include consultation, design and planning of research, execution, data analysis, reporting and advice, with a minimal turn-around time. In addition to behavioral observation, monitoring the physiological and emotional state of the consumer will be part of the service.

The techniques, methodology and service developed in the project can be used by the food and drink industry to (a) develop products for specific target groups (toddlers, teenagers, elderly people) that are hard to reach with traditional consumer research methods; (b) develop products for special occasions (party snacks, sports drinks, meal substitutes for elderly), or (c) check to what extent their general products are appreciated in different situations. The service can also be used by other organizations with an interest in food and eating and drinking behavior, including health aspects, such as hospitals, care centers, patient groups and insurance companies. It will contribute to a better understanding of eating and drinking habits and to ways to influence food choice behavior in a positive manner.

The service being offered enables the conduction of consumer behavior research on different locations under realistic conditions. Target groups that could not be reached before, can now be reached, which means that research can be done in more relevant target groups. Since food selection and consumption will be measured directly (and not indirectly via questionnaires or interviews), the objectivity and accuracy of the data will increase. Furthermore, the use of advanced computer software will make it possible to automate the measurement and analysis process to a great extent, allowing a drastic reduction in effort needed to execute studies and analyze results and thus a faster turn-around time of consumer testing.

Example: studying food consumption in a daycare center

An interesting setting to study eating behavior could, for instance, be a daycare center. Newborns, infants or toddlers can be observed while being fed or having their lunch at the center. Cameras and microphones will be set up in a room of the center. The setting will be recorded on video, with close-up view of the children's response to the food presented to them. After the session, caretakers are invited for a focus group discussion, which is recorded as well. This way, there is minimal disturbance of the normal daily activities at the daycare center.

The observed facial expressions, gestures, verbal and non-verbal behavior of the children will show what they like or dislike at what time of the day, to what extent they like variation, and in which frequency. The focus group discussions will reveal background information about the children and their eating habits. Strategies how to teach children to eat healthy food, and in what form, can be investigated.

Figure 3. *In the control room, the researchers can control the cameras, make video recordings, watch the videos, and annotate interesting behaviors using the latest observation software.*

In all tests, we will make sure that all activities are checked against privacy legislation to prevent any possible violation.

Project consortium

The project consortium consists of four partners with complementary assets and expertise:

- **Noldus Information Technology BV**, innovative developer of software, hardware and integrated solutions for research on human behavior, coordinates the project.
- **Centre for Innovative Consumer Studies (CICS)**, a leading research group of Wageningen University and Research Centre in the field of sensory science and food related consumer behavior. Together with Noldus Information Technology and two other partners, CICS has set up the Restaurant of the Future.
- **VicarVision BV** (Amsterdam), an R&D company in computer vision and the creator of FaceReader, a unique tool for the automated assessment of consumer emotions.
- **Symrise GmbH & Co KG**, a global supplier of fragrances, flavorings and active ingredients for the Fast Moving Consumer Goods industry – belonging to the top four in the international flavors and fragrances market.

References

1. Stewart-Knox, B., Mitchell, P. (2003). What separates the winners from the losers in new food product development? *Trends in Food Science & Technology*, **14**, 58-64.
2. Köster, E.P. (2003). The psychology of food choice: some often encountered fallacies. *Food Quality and Preference*, **14**, 359-373.
3. Köster, E.P. (2002). The specific characteristics of the sense of smell. In D. Dubois, C. Rouby & B. Schaal (eds.) *The specific characteristics of the sense of smell* (pp. 27-43). Cambridge University Press, Cambridge.
4. Van Herpen, E., Van Trijp, H., Kuipers, T. (2005). The influence of assortment organization on product comparisons and choice. *Proceedings of Measuring Behavior 2005, 5th International Conference on Methods and Techniques in Behavioral Research (Wageningen, 30 August – 2 September 2005),* 595-596.

Acknowledgements

Sponsors

The organizers of Measuring Behavior 2008 gratefully acknowledge the support from the following organizations:

Scientific Program Committee

- Harry Steinbusch, Conference Chair, Director School for Mental Health and Neuroscience, Director European Graduate School of Neuroscience; University of Maastricht, The Netherlands
- Andrew Spink, Scientific Program Committee Chair, Noldus Information Technology bv, Wageningen, The Netherlands
- Stephen Campbell, Faculty of Education, Simon Fraser University, Burnaby, Canada
- Nicky Clayton, Department of Experimental Psychology, University of Cambridge, UK
- Wim Crusio, Department of Psychiatry, University of Massachusetts Medical School, Worcester, MA, U.S.A.
- Marcel Dicke, Laboratory of Entomology, Wageningen University and Research Centre, The Netherlands
- Robert Gerlai, Department of Psychology, University of Toronto, Mississauga, Ontario, Canada
- Eco de Geus, Department of Biological Psychology, Vrije Universiteit, Amsterdam, The Netherlands
- Stan Gielen, Department of Medical Physics and Biophysics, Radboud University Nijmegen, The Netherlands
- Ilan Golani, Department of Zoology, Tel Aviv University, Israel
- Frans van der Helm, Man-Machine Systems & Control group, Department of Mechanical Engineering, Delft University of Technology, Delft, The Netherlands
- John Krantz, Department of Psychology, Hanover College, IN, U.S.A.
- Johan van Leeuwen, Experimental Zoology Group, Wageningen University and Research Centre, The Netherlands
- Gitte Lindgaard, Department of Psychology, Carleton University, Ottawa, Canada
- Josina Mojet, Agrotechnology & Food Innovations, Wageningen University and Research Centre, The Netherlands
- Lucas Noldus, Noldus Information Technology bv, Wageningen, The Netherlands
- Pertti Panula, Department of Biology, Åbo Akademi University, Turku / Abo, Finland
- Fabio Paterno, Laboratory on Human Interfaces in Information Systems, ISTI - C.N.R., Pisa, Italy
- Vicenç Quera, Department of Behavioral Science Methods, University of Barcelona, Barcelona, Spain

- Matthias Rauterberg, User-Centered Engineering, Faculty of Technology Management, Eindhoven University of Technology, Eindhoven, The Netherlands
- Berry Spruijt, Department Animal, Science and Society, Utrecht University, Utrecht, The Netherlands
- Franz Josef van der Staay, Department of Farm Animal Health, Faculty of Veterinary Medicine, Utrecht University, The Netherlands
- Hans van Trijp, Marketing and Consumer Behaviour Group, Wageningen University and Research Centre, Wageningen, The Netherlands
- Gerrit van der Veer, School of Computer Science, Open University Nederland, Heerlen, The Netherlands
- Louise Vet, Netherlands Institute of Ecology (NIOO-KNAW), Nieuwersluis, and Laboratory of Entomology, Wageningen University and Research Centre, Wageningen, The Netherlands

Local Organizing Committee

- Natasja Bogers, Chair
- René Bourgonje
- Leanne Loijens
- Lucas Noldus
- Gonny Smit
- Andrew Spink

Exhitibitors

Several companies operating in a broad range of behavioral-related disciplines will exhibit their products at *Measuring Behavior 2008*.

Animal Welfare

Animal Welfare is a young dynamic company, founded by biotechnical research analysts. We deliver products and customer based service especially aimed at biomedical research.
Our products are a true addition, for example:

- Speeding and improving the recovery of laboratory animals
- Improving animal welfare
- Easier work methods biotechnical/veterinary personnel

BASi

Contract research services and unique instruments for pharmaceutical industry. Featuring behavioral monitor to quantify laboratory animal behavior. Provides finer spatial and temporal resolution than laser-based devices, quantifies locomotor activity and tremor and seizure behavior. Replaces several pieces of equipment by performing functions simultaneously. For rats and mice between 15g and 500g.

Biopac

BIOPAC develops, manufactures, and supports data acquisition and analysis systems for life science research and education, including high-speed acquisition, amplifiers, telemetry, data loggers, stimulus presentation and virtual reality, transducers, electrodes and accessories. AcqKnowledge includes scoring and automation routines. BIOPAC is used in thousands of labs and cited in thousands of publications.

BIOSEB

BIOSEB develops equipments for the in Vivo preclinical behavioral research. Our latest instruments, including sensor automated FST, Home Cage ActiV-Meter for long term experiments, Weight-bearing tester for freely moving rodents , Temperature choice tests (T2CT and Gradient) will be presented. Come and see us for your instrument needs in behavior, pain, anxiety /depression...

Brain Products

Brain Products: A leading manufacturer of neurophysiological research products:

- Electrode caps for EEG & ERP
- Active Electrode caps
- Various EEG/ERP amplifiers up to 512 channels
- BrainAmp MR Amplifiers for use inside MRI scanner
- Analysis software including source analysis
- Education hard- and software

Compumedics Neuroscan

Compumedics Neuroscan is the world's leading developer and provider of research grade technologies for high-density EEG recordings, electro-magnetic source localization, multi-modal neuroimaging and enhancements to functional MRI. Neuroscan products are in use at over 1500 universities, corporate laboratories and national research institutes in approximately 40 countries worldwide.

Delta Phenomics
Delta Phenomics offers high-throughput phenotyping using PhenoMAP™, which allows objective collection and analysis of data in a novel paradigm: an enriched home cage provided with programmable stimuli, sensors and continuously video imaging. We transfer the experiment to the home cage, avoiding human intervention, animal handling and animal transportation.

Data Sciences International
DSI offers a wide variety of physiological monitoring solutions for respiratory, cardiovascular, and CNS applications involving acute or chronic studies in laboratory animals. Products include data collection and analysis systems coupled with hardwired amplifiers, implantable telemetry, and externally worn telemetry.BetterData, BetterScience.

Experimetria
Biomedical research and manufacture company. We produce devices mainly for pre-clinical investigations on fields of pharmacology & physiology.

MBRose
Equipment for feed intake and activity monitoring in home cage environments. We focus in complete solutions for mice, rats and minipigs, offering hygienic and easy to use home cage based platforms with a complete data collection system allowing the operators and researchers to efficiently monitor their experiments in real time without delay.

Metris
Metris develops and sells turnkey instruments for laboratory animal research.

The product portfolio includes LABORAS (recording and automatic recognition of rodent behaviors), SONOTRACK (recording and analysis of ultrasound vocalizations in animals), SleepSign (manual and automatic analysis of sleep stages in animals) and integration of behavioural data with several other parameters such as physiological data.

Major application areas are drug development, safety pharmacology, toxicology, pheno- typing.

NewBehavior
New Behavior combines state-of-the-art science with unique engineering. We provide customers not simply with hard- and software but also with support and advice encompassing the scientific use of our products in almost every possible field of research.

Noldus Information Technology
Noldus Information Technology aims to provide the best possible tools for behavioral research, allowing users to improve the quality of their data, increase productivity, and make optimal use of human or animal resources.

Our product range includes a number of successful products such as The Observer, EthoVision, FaceReader, CatWalk, and PhenoTyper . In addition, we offer custom software development, training, lab rental, and consulting services. Besides standard software packages, we offer integrated data acquisition and data analysis systems, including PCs and various sorts of audiovisual equipment, as well as complete observational laboratories.

Smart Eye

Smart Eye supplies the most innovative and cost efficient solutions in remote eye tracking. Smart Eye Pro represents the golden standard for head, eyelid and gaze tracking in the most difficult measurement situations.

SensoMotoric Instruments

SMI is a leading provider of eye tracking solutions for the scientific communities in fields including vestibular, ophthalmologic, psychological, reading, neurological, human factors, sports, ergonomic and marketing research. SMI´s "iVIEW" systems are applied in a broad range of environments from labs over driving and airplane simulators, shopping malls, virtual environments to cars and planes.

TEA S.A.S.

CAPTIV™: (1) Software for Task Analysis (2) Fystem for real-time acquisition, synchronization and processing of viedo, observations and physiological data: for EEG, EMG, ECG, skin conductance - including faceLAB™'s.

faceLAB™: 3D head, face, eye, eyelid and gaze tracking. Non-contact, with and without IR, automatic subject initialization, real-time data logging over network.

Ugo Basile

World leading manufacturers of instruments for physiological & pharmacological research, our design and production team is in close contact with the latest developments in biological research and draws on a wide variety of expertise. We manufacture instrumentation for pain assessment, motor coordination, ventilation, behavioral, activity, and metabolic monitoring.

Scientific Tours

IDEE (Engineering department, Maastricht University)

IDEE (Instrument Development, Engineering & Evaluation) is the engineering department of the faculty of Health, Medicine and Life Science (FHML) of the Maastricht University (UM) as well as for the academic hospital Maastricht (azM).

The goal of IDEE is to provide researchers and medical staff with innovative technology for scientific research as well as applied clinical research. IDEE specializes in design, development and prototyping of innovative instruments, devices, custom fit implants, software and equipment for a wide range of applications, with a focus on medical technology and life science.

Maastricht Instruments BV

Maastricht Instruments BV (MI) is an independent spin-off company (SME) residing under the Holding of University of Maastricht, and originates from IDEE, the engineering department of the FHML of the Maastricht University. MI and IDEE closely cooperate, where MI is the partner for all external relationships and business relationships, whereas IDEE performs activities for the FHML and azM.

Prototypes that have been developed by IDEE for research purposes are redesigned by MI to fit market specifications, depending on the needs of the market. MI also performs engineering tasks for external companies and institutes.

Visitors of the tour will experience the engineering capabilities of IDEE and MI and the specific expertise in the field of signal and image acquisition and analysis by demonstrating the following equipment:

- MPAQ system, a versatile high-end multi channel data-acquisition system
- PASAQ system, a ambulatory wireless data-acquisition system
- Fall & Balance detector
- VET, a Video Eye Tracking Device for vestibular research

The tour will be completed with a visit at the vestibular research lab of Prof. Kingma, located at the azM.

Simulator Lab (Work and Organizational Psychology, Maastricht University)

Work occupies a central place in the lives of many people, not only because a large part of one's life is spent at work, but also because of the meaning that people derive from their work. Work and organizational psychology studies work as an activity, from the perspective of individual workers as well as the organization.

For the purpose of studying aviation related (group) behavior, our simulator lab is equipped with a static real-life helicopter cockpit environment simulating a Robinson 22/44 training helicopter. The two-seat simulator provides a human machine interaction platform for research into single pilot behavior, and two-member team behavior in an instructional setting, regarding navigation and way-finding, decision making, situational awareness and stress. Helicopter flight characteristics are simulated in slow and normal flight with a real time interface to individually programmable dials and controls. It is possible to manipulate the timing of dial readouts lags, and of emergencies such as weather deterioration and engine malfunctions. Computer monitored flight behavior, optionally with physiological measures and camera observation are measured.

Academic Anxiety Center (Clinical and Experimental Psychiatry, Department of Psychiatry & Neuropsychology, Maastricht University)

The research group 'Clinical and Experimental Psychiatry' is a subdivision of the department of Psychiatry & Neuropsychology at Maastricht University. This group has conducted internationally leading research in the field of anxiety (disorders) for over 20 years. The backgrounds of the different team members are diverse (psychiatrists, psychologists, biologists) underwriting a broad scope. Furthermore, there is a close collaboration with the Academic Anxiety Center (Mondriaan Zorggroep, Maastricht); a highly specialized treatment facility. Prof. dr. E. Griez is chair of the research group as well as one of the founders of the Anxiety Center.

The research group is internationally recognized for experimental modeling of anxiety in humans, and in particular for the development of the panicogenic 35% CO_2 inhalation. The specific panicogenic characteristics of this model allow researchers to study various aspects of panic under standardized laboratory circumstances. Key topics include: treatment effects, biological mechanisms, genetics and dependence. The tour will provide you with the opportunity to learn all about the 35% CO_2 inhalation methodology. Equipment will be presented, as well as the measures that are currently in use (electronic questionnaires) and future perspectives (fully integrated physiological measures). Also, for those who are interested: it is possible to experience a 35% CO_2 /65% O_2 inhalation yourself.

Institute of Experimental Psychopathology (Department of Clinical Psychological Science, Maastricht University)

The Institute of Experimental Psychopathology investigates psychological processes and mechanisms in human psychopathology, preferentially using experimental study designs. During the tour experimental set-ups will be demonstrated that measure cognitive and genital responses to erotic stimuli and behavior in response to pain induction.

Among the techniques that will be shown are male and female genital plethysmography, and computer-based reaction-time tests to measure speeded detection of and slowed disengagement from specific stimuli (visual search task). The set-ups that enable measurement of pain-related behavior include equipment to (safely) deliver electrical shock, pain due to (finger) pressure, and pressure applied to the vaginal entry by an inflatable balloon-device.

Visitors of the tour can experience the use of an ‘actometer’ that measures movement in real-life situations.

Human Energetics Facilities (Department of Human Biology, Maastricht University)

At the Department of Human Biology we conduct high standing fundamental research into mechanisms of energy balance. The main focus is on intervention strategies for long-term maintenance of energy balance. Parameters of study are (neuro) endocrine mechanisms and the interaction with environment and behavioral components including food intake and physical activity.

A food intake research lab is used, including a universal eating monitor, respiration chambers, and ventilated hood devices. Food reward is assessed by means of quantifying wanting and liking. Compliance to energy intake is measured using the double labelled water method for determining energy expenditure. Genetic predisposition is measured by determining relevant polymorphisms for fat cell differentiation, body composition, muscle fibre type distribution, behavioral disinhibition, dietary restraint. Body composition is assessed using underwater weighing, deuterium dilution, DEXA and MRI. Physical activity and activity patterns are assessed using tri-axial accelerometry.

Visitors of the tour can see:

- Laboratory restaurant with universal eating monitor
- Underwater-weighing and Bod-pod for body composition
- Respiration chambers
- Ventilated hoods
- Mass spectrometry facilities for labelled water to assess total energy expenditure, water turn over and water balance
- Physical activity assessment in daily life with miniature tri-axial accelerometers

Ambulatory Motion Analysis (Department of Human Movement Sciences, Maastricht University)

At the department of Human Movement Sciences we focus on understanding how skeletal muscle weakness due to aging and/or chronic diseases affects movement performance, mobility, physical activity and thereby health. The department of Human Movement Sciences has strong collaborations with clinicians (gerontology, internal medicine, pulmonology, rehabilitation, orthopedics) in projects aimed at transferring ‘state of the art’ laboratory knowledge on human movement and muscle function into simpler measuring instruments for functional assessment outside the laboratory environment.

The department has a fully equipped 3-D motion analysis laboratory (VICON-NEXUS) and muscle testing facilities (Biodex) as well as a wide range of ambulatory motion analysis and ExG tools for long term monitoring of physical activity and physiological parameters. Recently a wearable fall detector/activity monitor for use in nursing homes and care centers has been developed in close collaboration with Maastricht Instruments BV.

Vistors of the tour will see:

- 3D gait analysis
- Fall and balance detector
- Ambulatory system for activity monitoring of upper and lower limbs
- Sensorsleeve for arm function monitoring in MS patients

Author Index