

The Science Behind Measuring Behavior

P. Pohl

Child Psychology Practice Garmisch, Garmisch-Partenkirchen, Germany. dr.pohl@kinderpsych-garmisch.de

A major development in 20th Century psychology and behavioral research was the adoption of the methods and statistical analyses of the English statistician, geneticist and evolutionary biologist R.A. Fisher (1890-1962) as the almost exclusive model of research practice and data analysis. The rapidity and thoroughness of this movement was remarkable as it comprehensively re-defined, within a period of only about 30 years (1925-1955), what „science“ was, is, and ought to be in the field of psychological and behavioral research. Spreading from applied to experimental research, this paradigm gave the world factorial designs, randomization of research elements to treatment conditions, the analysis of variance (ANOVA), the analysis of covariance (ANCOVA), the null hypothesis, and null hypothesis tests of significance (NHST). Fisher’s statistical tests were not the only practices adopted. His advocacy of factorial designs, i.e., experiments investigating more than one level of an independent variable, was equally influential, reflecting his view that experimental design and statistical analysis are only “different aspects of the same whole” [1]. Beginning in the mid-1950s and continuing up to the present, more than 90% of the articles published in psychological and behavioral research typically report the outcomes of significance tests [2], thereby demonstrating that the preoccupation with inferential statistics was and still is the main methodology in psychological and behavioral research with the underlying assumption that scientific knowledge emerges out of data through the application of statistical analysis.

Clearly, reproducible group data describe some kind of order and may well form the basis of a science. It cannot, however, be a science of individual behavior nor can it be a science of group behavior. It is, in fact, a science of averaged behavior of individuals linked together only by the averaging process itself [3]. Resulting from this state of affairs, disenchantment over the prevailing statistical methodology and its questionable relevance for understanding the behavior of the individual has led to a debate over the applications of statistics in psychological research, culminating in the establishment of a Task Force on Statistical Inference (TFSI) by the American Psychological Association (APA) in 1999. However, conspicuously lacking in the TFSI report is any recognition that single-case research designs might be an alternative to current statistical practices including statistical pattern recognition. This is curious because single-case research is based on a rejection of null hypothesis testing, emphasizes the visual analysis of data in graphs, and relies on replication as the key to making reliable causal inferences. Presumably the oversight is partly a result of the persisting confusion about the differences between case studies and single-case designs. Whereas case studies are usually clinical descriptions of individual cases and normally cannot support valid causal inferences, single-case research designs are controlled experiments from which valid causal inferences may be drawn. This neglect has deprived many researchers and practitioners in many areas of psychology and behavioral measurement of a powerful methodology for evaluating the effects of various psychological interventions and experiments. Among the numerous benefits of using single-case research methodology are that it helps researchers and practitioners to focus on the individual (which is a widespread necessity in clinical and applied psychology), reduces some of the problems inherent in averaging across groups, makes it easier to undertake scientific investigations of rare and unusual cases and phenomena, and facilitates ethical innovation and professional accountability. At the very least, turning towards single-case methodology would help behavioral researchers to eliminate the „double standard“ in their treatment of individuals, i.e. claiming to be able to apply science to individual organisms while working almost exclusively with group-aggregate data.

Why, then, has single-case methodology been used so rarely in behavioral research in the past decades? An analysis of 40 English-language introductory psychology textbooks published between 1979 and 2009 provides a clue: Investigators Hobbs and Chiesa summarized that „the history of psychology in the twentieth century as presented in introductory textbooks works to the detriment of behavior analysis and to the disadvantage of those students and professionals who would welcome and use behavior analysis if they were properly informed“[4]. In

this context, it is important to mention that by about the middle of the 1950s, just as the so-called „cognitive revolution“ began to gather momentum, behavior analysis and the single-case research methodology it had engendered, soon fell into disrepute and has been, for the most part, ignored by mainstream psychology to this day. It is no exaggeration to say that the issues surrounding this curious circumstance are fundamental to the science of behavior and will be dealt with in detail in the conference presentation.

And finally, based on the methodological argument developed here, human behavior studies employing *Observer XT* [5] and *FaceReader* [6] and animal studies employing *EthoVision XT* [7] and *PhenoTyper* [8] will be discussed with a view towards exemplifying the advantages of single-case methodology in measuring behavior.

References

1. Fisher, R.A. (1935). *The design of experiments*. London, UK: Oliver & Boyd, p. 3.
2. Hubbard, R., Parsa, R.A., & Luthy, M.R. (1997). The spread of statistical testing in psychology. *Theory & Psychology*, **7**, 545 – 554.
3. Sidman, M.(1960). *Tactics of scientific research*. New York: Basic Books.
4. Hobbs, S. & Chiesa, M. (2011). The myth of the „cognitive revolution“. *European Journal of Behaviour Analysis*, **12**, 385-394, p. 385.
5. Reissland, N. & Stephenson, T. (1999). Turn-taking in early vocal interaction: a comparison of premature and term infants' vocal interaction with their mothers. *Child: Care, Health, and Development*, **25** (6), 447-456.
6. Wei He, Boesveldt, S., de Graaf, C. & de Wijk, R. (2014). Dynamics of autonomic nervous system responses and facial expressions to odors. *Frontiers in Psychology*, **5**, 1-8.
7. Spink, A.J., Tegelenbosch, R.A.J., Buma, M.O.S. & Noldus, L.P.J.J. (2001). The EthoVision video tracking system – A tool for behavioral phenotyping of transgenic mice. *Physiology & Behavior*, **73**,731-744.
8. Brielmaier, J., Matteson, P.G., Silverman, J.L., Senerth, J.M., Kelly, S., Genestine, M., Millonig, J.H., DiCicco-Bloom, E., & Crawley, J.N. (2012). Autism-relevant social abnormalities and cognitive deficits in *Engrailed-2* knockout mice. *PLoS ONE*, **7**, e40914.