

Assessing the Personality Trait Compliance in a Game Context

Christof van Nimwegen¹, Herre van Oostendorp¹, Alec Serlie², Joost Modderman²

¹*Department of Information and Computing Sciences, Faculty of Science, Utrecht University, Utrecht, The Netherlands.
C.vanNimwegen@uu.nl*

²*GITP Research, Rotterdam, The Netherlands. A.Serlie@gitp.nl*

Introduction

Serious games for contexts *not primarily aimed at entertainment* are a well-established phenomenon. Well-designed serious games captivate players for a specific purpose, and in spite of the often non-entertainment purpose they provide engaging environments. Players are provoked to endeavor in higher-order cognitive skills such as handling of information, decision making and solving problems. Serious games are widely implemented in areas as diverse as healthcare, the military, and persuasive technology or applications aimed at attitude change [1]. There is also the rising trend to use serious games in recruitment contexts, e.g. to acquaint people with certain professional areas and enthuse them, of which America's Army [2] is a famous example. Using serious games for *assessment* in a recruitment procedure, is one step further, which is gaining popularity as a screening means next to standard interviews. Playing a game with the goal of assessment is about uncovering personal traits, skills and capabilities (or, of course, deficiencies) that are present in persons. Personality tests have been in use for a long time already e.g. for employee selection, team building or training. They can be paper tests, interviews, or computerized versions of these, but there are a few disadvantages, one of them being the fact that it is sometimes too easy to give socially desirable answers. Furthermore, a self-report questionnaire does not provide information about the players' decision process. Our idea is that accommodating assessment in a game (although not purely meant to entertain) provides a richer, less repetitive and predictable, and more immersive context. Players might act more natural and intuitive compared to holding a pen and paper, and thus contribute to the validity of the measurement. Serious games can uncover individual personality traits of people in a structural manner, uncolored by the chemistry between interviewer and (future) employee.

The Compliance Game

This paper describes the process of developing a first version of "The Compliance Game" and its evaluation, followed by a second version and its evaluation. An additional aim of the project, using the game as a test bed, is to develop a notation and visualization method called "GameDNA" (Game Discourse Notation and Analysis). It is player-centric and accommodates mental but also other actions that players face in the game, along with the discourse between system and player [3]. In serious games where balancing between mental decisions and effort from the player, and reactions to the system is the issue, this becomes important. With this tool people from different disciplines (as in this research project) can clearly see what the player has to do, how the flow of information is shaped and how the prospected measurable features are balanced and correctly implemented.

There were three partners in the project. GITP is an international Human Resource Management consultancy firm whose goal is to enhance organizational effectiveness and performance by assessing people. In this project, they wanted a game developed to measure "Compliance": the tendency of an individual to agree with an instruction from another person, even if he/she does not believe it, only to satisfy the other(s) or to avoid conflicts. Serious games have been used in the context of compliance, such as "Straight Shooter!" by [4], a real-time 3D e-learning game aimed at compliance. However, this involved training, while our aim was assessment, uncovering which attitudes are already present in people *before* they possibly enter a company. Compliance is a hot topic today in the aftermath of financial crises in banking and financial services. There is a need for better assessment, as the rules are getting stricter and policies are changing. The technical development was done by Ranj, a serious game studio that is endeavoring to serious games for assessment, such as their award winning "Houthoff Buruma The Game" [5]. The third partner is the Department of Information & Computing Sciences (ICS) at Utrecht University and is involved in various gaming projects such as GATE [6].

The Compliance Game version 1

Content

In the Compliance game, the player acts in a setting in which Compliance and its associated behavior is an everyday occurrence: the sales department of an international company that manufactures smart connected devices (e.g. computer pads). The player acts in a story, makes decisions and expresses opinions. In the scenes in the game assignments must be carried out (or not), managerial decisions must be taken, suggestions to act immoral (but profitable for the company) must be acted upon.

The game “storyline” was made as engaging as possible, and was constructed such (especially the dialogues) that players will believe that the game characters reactions are a specific reaction to the player’s own actions. This is not the case in reality, the game is sequential in the sense that all players have to pass the same milestones, although there is of course freedom in actions/answers during the game. The measurements (see next paragraph) are quantitatively analyzed afterwards, therefore it should be safeguarded that all players are confronted with the same events to act on. The building stones of the game are:

- Animated realistic characters (Figure 2) which are moving and talking (pictures of humans shown as moving stills, audio fully streamed). They talk to the player and ask questions.
- Text shown, e.g. the actions the player can choose from.
- A Smartphone (Figure 1). Text messages, emails and voice message arrive here, and can be read/listened to.
- Actions: grab phone with the mouse, open an item and close it, and choose actions and make decisions by clicking on one of them.

Measurements in the game

What is measured is the construct “Compliance” and 4 constructs of “Machiavellianism”, one’s propensity to distrust others, engage in amoral manipulation, seek control over others, and seek status, for which scales already exist in reliable reference tests [7]. 5 constructs are measured: Compliance, and of Machiavellianism: “Distrust in others”, “Amorality”, “Desire for control” and “Desire for status”. The actions in the story that players choose, correspond to a 4-point Likert-scale of which the answers range on the spectrum on one of these 5 constructs. A typical situation in the game is that the player is confronted with the dilemma whether to accept assistance from someone else or not (there might be a hidden agenda behind it), involving considerations of whether and how to delegate. The player has to indicate the degree to which he/she would act compliant on a certain proposition (Figure 3).

What this game should accomplish is to measure the same as the reference tests do, i.e. correlate highly. If



Figure 1. Using the smartphone in the game.



Figure 2. A talking and moving character in the game.

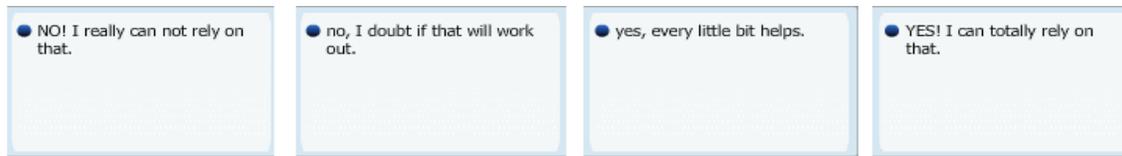


Figure 3. Answers in the game players can choose from ranging on a (in this case “distrust”) scale

similar scores on the 5 constructs are yielded, one can say that the game can be used in a valid way to measure such personality traits. The first game version consisted of 5 scenes. Each scene is a part of the story and accommodates 3 different constructs in the storyline on which the player has to act. In each scene, there are 3 actions pertaining to the constructs involved, 2 actions during dialogue, and 1 action at the end. When the game is played, there are 3 measurements on each construct, which are subsequently averaged to a score indicating the position of a player on that construct.

Evaluation of game version 1

The first game version was played by 72 test participants from diverse age groups, backgrounds and professions. After playing the game, they also filled out the “classical”, validated reference test (online) consisting of 36 questions, aimed at measuring the same constructs. This in total took about 45 minutes. The aim in this phase of the project was to have a game that yields the similar scores as the reference tests. Although at first glance similar *average* scores on the constructs of the game and reference tests were found, the results were not yet convincing. On none of the constructs, correlations between game score and reference test score reached significance. Analysis showed that the problem probably was the reliability; the internal consistency (Cronbach's alpha) of the 3 items of the 5 constructs was very low, respectively Compliance .30, Distrust .35, Amorality -.32, Control .15 and Status .14. In other words, the separate items of each construct did not quite measure the same thing. Qualitative questionnaires that were administered in addition revealed that in general, the test participants claimed to like to play this game, at times they felt “immersed” and found it more interesting (and at times more difficult) than filling out questionnaires.

The Compliance Game version 2

Development

After the informative, but not sufficiently satisfying results of the evaluations of the first game version, careful analysis took place. Using GameDNA, the game and the empirical results were carefully dissected and we scrutinized the places where inconsistent answers took place. The measurements of each construct were revised or replaced, and the amount of measurements was extended to 4, in order to obtain more reliable means that are based on more scores. The game was shortened a bit, answers that had the risk of ambiguity were refined, unnecessary interaction between system and player removed and more details were changed, including some graphics and audio.

Evaluation

Now there were 64 participants from diverse age groups, backgrounds and professions. The preliminary results are slightly more promising. Amorality in the game and amorality by the reference test had a significant correlation, $r(63) = .53, p < .01$, and Cronbach's alpha of this construct (4 items) measured in the game was in the right direction at .42. Compliance in the game and compliance by the reference test also had a significant correlation, $r(63) = .25, p < .05$., but surprisingly, Cronbach's alpha of the 4 items of the construct was '0'. The correlations between game scores and reference test scores on “Distrust”, “Desire for control” and “Desire for status” still did not reach significance yet. Cronbach's alpha of these constructs were low; .22, .12 and .18 respectively. Summarizing, on 2 of the 5 constructs similar scores were obtained both in the game and in the reference test.

Conclusion and Discussion

Stakeholders from different disciplines cooperated on our assessment game measuring Compliance by embedding psychological constructs in a serious game. We have seen that it is very difficult to consistently replicate personality trait constructs as the one we envision. The first game version did not result in convincing correlations between game scores and reference test scores. But since all moves and actions of the players were logged, the sessions were highly informative because by using GameDNA [3] we could see where it went wrong in the game. Unlike when measuring personality traits with standard non-game reference tests, the player acts in a story. The questions to be answered are not standalone; story elements of a scene can yield a situated judgment of one kind on a construct, while in another scene the judgment on a similar question is different. Careful iteration, using GameDNA does on the one hand dissect the game events and actions, and on the other hand it provides a helicopter view of the discourse and information flow in the game story. This resulted in a second, more balanced and elegant game version. With the second game version on 2 of the 5 constructs similar scores were obtained both in the game and in the reference test. For yet another iteration, it is useful to scrutinize on “Distrust”, “Desire for control” and “Desire for status”. An iterative approach as described here might prove imperative in the development and validation of serious games of the type we envision. The multidisciplinary nature of the project, cooperating with domain experts in assessment and game developers, calls for one shared language to shape the narrative and map game events and information flow. For this purpose, GameDNA will be further developed and applied in the course of future game development.

References

1. Susi, T., Johannesson, M., Backlund, P. (2007). Serious games - an overview. Technical Report HS-IKI -TR-07-001. University of Skövde.
2. America's Army[®] (2002). <http://www.americasarmy.com>. Accessed 20 February 2012.
3. Nimwegen, van C., Oostendorp, van H., Modderman, J., Bas, M. (2011). A test case for GameDNA: conceptualizing a serious game to measure personality traits. Proceedings of the *16th International conference on Computer Games: AI, Animation, Mobile, Interactive Multimedia, Educational & Serious Games* (CGAMES 2011 USA), IEEE digital library.
4. Prensky, M. (2004). *Digital Game-Based Learning*. New York: McGraw-Hill
5. Houthoff Buruma. The Game[®]. <http://www.experiencehouthoff.com>. Accessed 2 January 2011.
6. GATE. <http://gate.gameresearch.nl/>. Accessed 15 December 2010.
7. Dahling, J.J., Whitaker, B.D., Levy, P.E. (2009). The development and validation of a new Machiavellianism scale. *Journal of Management* **35**, 219-257.