**Gamified and “Classical” Psychometric Assessment – Convergent, Divergent, and Incremental Validity**

Katharina Lochner1, Lara Montefiori2, Achim Preuss1, & Maria Panagiotidi3

1 cut-e Group

2 Arctic Shores

2 Staffordshire University

Ability tests and personality questionnaires have been shown to be valid predictors of job performance in numerous studies (e.g., Robertson & Smith, 2001; Schmidt & Hunter, 1998). However, research also shows that social acceptance of these types of instruments is low compared to interviews or work samples (Hausknecht, Day, & Thomas, 2004) and that consequently candidates do not enjoy them. Game-based assessment (GBA) strives to fill this gap by building on the fact that humans perceive games engaging and enjoyable (Hamari, Koivisto, & Sarsa, 2014). However, the question is whether this kind of assessment taps into the same constructs as “classical” assessments. A few studies indicate that this is true for different personality traits (DiCerbo, 2014; Ventura, Shute & Zhao, 2013; Yee, Ducheneaut, Nelson & Likarish, 2011; Whon & Wash, 2013) and feasible for ability (van Lankveld, Spronck, van den Herik & Arntz, 2011).

The purpose of this study was to explore to what extent similar constructs measured by tests and questionnaires on the one hand and by the GBA “Cosmic Cadet” correlate. Hypotheses are as follows:

Cognitive abilities: Positive correlations between…

* Processing Capacity and fx (H1).
* Processing Capacity and lct (H2).
* Innovation Potential and fx (H3).
* Innovation Potential and lct (H4).
* Innovation Potential and fluency, flexibility, and originality (H5).
* Executive Function and fx (H5).
* Executive Function and lct (H6).

Personality traits:

* Positive correlation between Persistence and energetic (H7).
* Negative correlation between Risk Appetite and conscientious (H8).

**Method**

**Participants**

*N* = 148 participants (56.8% male) were recruited through Prolific Academic, a portal that allows participants to take part in online studies for small monetary rewards. The mean age was 23.6 (*SD* = 5.9, range = 16 – 51). Only participants who reported their first language as English were allowed to take part in the study.

**Procedure**

An online study was posted on Prolific Academic inviting participants to take part in a study, which involved downloading and playing a game based assessment and completing a number of cognitive ability tests. Participants were reimbursed for their time.

**Measures**

**Cosmic Cadet.** Cosmic Cadet is a GBA in which candidates take the role of astronauts that have to collect stars in outer space. There are several levels, each of which has a different task to complete. The game assesses cognitive abilities such as processing capacity and personality traits such as resilience. Table 1 provides an overview of the characteristics assessed by the game.

----------------------------------------------------

Insert Table 1 about here.

----------------------------------------------------

The game has a test-retest reliability of *r =* .84. Table 2 depicts information on validity.

----------------------------------------------------

Insert Table 2 about here.

----------------------------------------------------

**scales fx.** The cognitive ability test measures deductive logical reasoning. The test has a split-half reliability of .84, a test-retest reliability of .81, and a correlation of *r* = .34, *p* < .01 with scales cls (cut-e Group, 2004), a test measuring inductive-logical reasoning.

**scales lct.** The cognitive ability test measures the ability to memorize objects shown for a split second in their correct order. The test is made up of several sections and thus it examines the learning aptitude of a candidate as well. It has a split-half reliability of .88 and a correlation of *r* = .32, *p* < .01 with Raven’s Advanced Progressive Matrices (Raven, Raven, & Court, 2003).

**sparks.** The cognitive ability test measures creativity and its three subscores fluency, flexibility, and originality. Test-retest reliabilities of the scores are: fluency: .82; flexibility: .67; originality: .71 (retest interval: two weeks). Concurrent validity with Torrance Tests of Creative Thinking (TTCT; Torrance, 1974) is: fluency: .99; flexibility: .85; originality: .88.

**shapes.** The questionnaire is an adaptive computerised questionnaire and provides a detailed assessment of a person’s job-related personality traits. Table 3 provides an overview of the traits measured, Table 4 depicts internal consistencies and test-retest reliability.

----------------------------------------------------

Insert Tables 3 and 4 about here.

----------------------------------------------------

**Results**

Table 5 provides an overview of correlations between scores from the ability and personality assessments on the one hand and the scores from Cosmic Cadets on the other hand.

----------------------------------------------------

Insert Table 5 about here.

----------------------------------------------------

**Discussion**

The purpose of this study had been to explore the overlaps between cognitive ability and personality traits measured by different methods, with GBA on the one hand and ability tests and personality questionnaires on the other. The results show that there is considerable overlap within convergent thinking (Processing Capacity and Processing Speed on the one hand and fx and lct on the other). Convergent thinking (fx and lct) is also related to Innovation Potential.

However, counter to the hypotheses, there is no correlation between Innovation Potential and fluency, flexibility, and originality as measured by the creativity test sparks. This could be due to the fact that Innovation Potential has a wider definition of innovation in the sense of not only coming up with new ideas, but also seeing where new ideas might be needed and putting them into practice, whereas sparks only measures creativity in a narrower sense.

As expected, there was a negative correlation between Risk Appetite and conscientious, however, counter to expectations there was no significant correlation between Persistence and energetic. Again, this might be due to the fact that constructs measured are slightly different. Moreover, a few correlations that had not been hypothesized showed up and will need to be explored further. In summary, it seems that Cosmic Cadets measures constructs that are related to the classical psychometric assessments, but that cannot be fully explained by the latter. This could be seen as a problem, but also as an opportunity because the game taps into something that cannot be measured by the classical instruments.

One shortcoming of this study is that no job performance criteria were included. Based on the previous research and based on the reliability and construct validity data it is likely that Cosmic Cadets has incremental validity with respect to job performance, and it is worthwhile exploring this further in future studies.

References

cut-e Group (2004). *scales cls. Quality Characteristics*. Hamburg: cut-e Group.

DiCerbo, K. E. (2014). Game-based assessment of persistence. *Journal of Educational Technology & Society*, 17(1), 17-28.

Hamari, J., Koivisto, J., & Sarsa, H. (2014, January). *Does Gamification Work? – A Literature Review of Empirical Studies on Gamification.* 47th Hawaii International Conference on System Sciences (HICSS), Waikoloa, Hawaii.

Hausknecht, J. P., Day, D. V., & Thomas, S. C. (2004). Applicant Reactions to Selection Procedures: An Updated Model and Meta-Analysis. *Personnel Psychology, 57*, 639–683. doi:10.1111/j.1744-6570.2004.00003.x

Schmidt, F. L., & Hunter, J. E. (1998). The Validity and Utility of Selection Methods in Personnel Psychology: Practical and Theoretical Implications of 85 Years of Research Findings. *Psychological Bulletin, 124*(2), 262–274.

Raven, J., Raven, J. C., & Court, J. H. (2003). *Manual for Raven's Progressive Matrices and Vocabulary Scales*. San Antonio, TX: Harcourt Assessment.

Robertson, I. & Smith, M. (2001). Personnel Selection. *Journal of Occupational and Organizational Psychology, 74*, 441-472.

Torrance, E. P. (1974). *Torrance Tests of Creative Thinking*. Scholastic Testing Service, Inc.

van Lankveld, G., Spronck, P., van den Herik, J., & Arntz, A. (2011). Games as personality profiling tools. In *Computational Intelligence and Games (CIG), 2011 IEEE Conference on* (pp. 197-202). IEEE.

Ventura, M., Shute, V., & Zhao, W. (2013). The relationship between video game use and a performance-based measure of persistence. *Computers & Education, 60*(1), 52-58.

Yee, N., Ducheneaut, N., Nelson, L., & Likarish, P. (2011, May). Introverted elves & conscientious gnomes: the expression of personality in world of warcraft. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 753-762). ACM.

Wohn, D. Y., & Wash, R. (2013). A virtual “Room” with a cue: Detecting personality through spatial customization in a city simulation game. *Computers in Human Behavior, 29*(1), 155-159.