The Effects of Adding Premise and Backstory to Psychological Tasks

Max V. Birk
University of Saskatchewan,
Department of Computer Science
Interaction Lab
max.birk@usask.ca

Benjamin Buttlar
University of Trier
Department of Psychology
s1bebutt@uni-trier.de

Regan L. Mandryk
University of Saskatchewan,
Department of Computer Science
Interaction Lab
regan@cs.usask.ca

Jason Bowey
University of Saskatchewan,
Department of Computer Science
Interaction Lab
jtb134@mail.usask.ca

Abstract
Psychological tasks are used for assessment, induction, and treatment in both research and therapeutic contexts. Adding game-elements (e.g., leaderboards, premise, or points) could be beneficial for participant motivation; however, it has been argued that the value of gamification could come at a cost to experience or reliability of the task. We replicated four psychological tasks and added premise and backstory to investigate the effect of gamification on task performance and player experience. Our results show that adding game elements has the potential to negatively influence both player experience and performance.

Author Keywords
gamification; psychology; performance task; cognition

Introduction
Psychological tasks are used in both research and applied psychology (e.g., educational, clinical, or industrial) for a variety of reasons, including: to measure performance, induce a psychological state, or assess and treat psychological issues. To allow for better standardization as well as cheaper deployment and analysis, the presentation of these tasks is often computer-based. As such, designers have considered leveraging the power of gamification to motivate participation in standard psychological tasks.

Participants
218 participants (46.3% female) with an average age of 32.79 (*SD=10.29*) participated. Participants received $6 compensation paid through the platform.

Platform
Amazon Mechanical Turk (MTurk), a platform that acts as a broker between parties offering a range of *Human Intelligence Tasks* (HITs) (e.g., marketing questionnaires, or research studies) and paid workers.

Player Experience
PENS: Measures *competence*, e.g. experiencing success and failure based on one's own skills, *autonomy*, e.g. accepting challenge under one's own volition, and *relatedness*, e.g. experiencing relations to others.

IMI: Measures *enjoyment*, e.g. "I enjoyed this game very much", *tension*, e.g. "I felt tense while playing the game", and *effort*, e.g. "I put a lot of effort into this game."

PANAS: Measures positive affect and negative affect.

Adding game-elements in a non-game context has been shown to be effective in a variety of settings (e.g., the ESP game [9], Foldit [5]). In the context of gamifying standard psychological tasks, research has shown that game elements can help to increase task accessibility [3], foster engagement with a system and motivate participation in a treatment over the long-term [6]. As intriguing as these benefits are, research has yet to show whether or not they come at a cost, such as decreased reliability or changes in performance.

Adding game-elements into computer-based psychological tasks for the purposes of assessment, induction, or treatment has several ethical and practical implications. First, a gamified task can deviate in terms of the user experience; however, the *reliability* of assessment or treatment needs to remain equal to or better than the un-gamified task. Second, gamified tasks can be embedded into games and could be used to assess or treat users without their *explicit consent*. Third, the *increased motivation* of players of a gamified task has implications for the interpretation of findings.

To begin to understand how task performance and player motivation differs in gamified and standard psychological tasks, we investigated the effects of *adding premise* to four psychological tasks – Go/No-Go, N-Back, ambiguous word interpretation, and facial feedback. We focused on adding premise as our gamification element as it would not change the nature of the standardized task through changing the mechanics, feedback, or reward structure. In addition, the use of exposition is one of the tools recommended for meaningful gamification of systems [7].

Our results show that adding game elements has the potential to negatively influence player experience and performance, demanding careful evaluation before being applied in research or therapeutic environments.

Methods

Procedure
After participants gave consent they played one of the games and filled out player experience measures (see sidebar) after each task.

Tasks with and without premise
The four different psychological tasks were presented to half of the participants as a task, and to the other half as a game with premise added. We kept all other elements of the task stable, and only added premise through graphical assets and a backstory (reasoning) presented prior to the task. We used a Zombie theme, which allowed us to present a premise for different scenarios, e.g., shooting, running away, selecting.

Go/No-Go Task: To test executive functioning we used a Go/No-Go Task [7]. The task presents a sequence of stimuli for 500ms. In the task, participants respond to circles, but not to squares; in the game condition, players shoot blond zombies, but not a mole with a yellow hat. Higher precision and accuracy indicate better functioning.

N-Back: For each stimulus in the N-back task [4], participants are asked to indicate if the stimulus is the

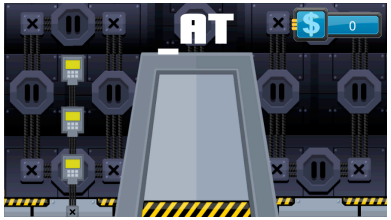

Figure 1 The gamified version of the ambiguous word task.

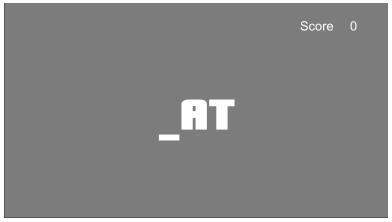

Figure 2 The standard version of the ambiguous word task

Figure 3 The gamified version of the N-back task

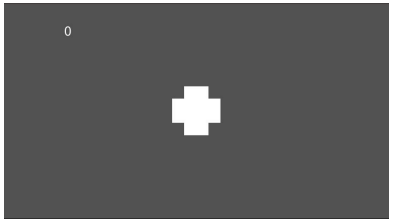

Figure 4 The task version of the N-back task

same as the one presented 2-back (press E), or not (press O). The stimuli were letters in the task version and zombies in the game. Higher precision and accuracy indicate better short-term memory.

Facial Recognition: In this task [2], a neutral face is presented and changes for 500ms to an emotion (happy, sad, angry, surprised), which participants must identify through a key press. The game was presented as a school for zombies to learn to recognize human emotion. The sum of correctly-identified expressions is a measure of emotion recognition.

Ambiguous Word Task: The ambiguous word task [1] presents words that can be interpreted as neutral or aggressive, e.g. S_AY (SLAY or STAY). In the game, participants were told to escape from a warehouse by guessing the password to open doors. Performance is the sum of hostile words chosen.

Results

Go/No-Go Task: We found no differences in player experience. However, the sensitivity was higher for the task *($F_{1,40}$=5.74, p<.021, η^2=0.126).*

N-Back: We found significant differences neither for player experience measures, nor for task performance.

Facial Recognition: Players experienced higher relatedness *($F_{1,59}$=4.74, p=.03, η^2=0.07)* in the plain task. Performance did not statistically differ.

Ambiguous Word Task: Players experience higher levels of enjoyment *($F_{1,68}$=5.62, p=.021, η^2=0.08)*, autonomy *($F_{1,68}$=17.02, p<.001 η^2=0.20)*, relatedness *($F_{1,68}$=9.80, p<.01, η^2=0.126)*, and immersion *($F_{1,68}$=8.05, p<.01 η^2=0.106)* in the plain task.

The ratio of hostile words to words overall differed *($F_{1,68}$=4.00, p<.049, η^2=0.056),* revealing that the game increases aggressive interpretation of words.

Interpretation

Our results show that the gamification of the N-Back task and the Facial Recognition task do not show differences in performance measures; however, gamification of the Go/No-Go task decreased performance in a strong effect. This difference is likely a result of the discrepancy of the stimuli – in the task, a circle is very different from a square, while the complexity of the more detailed characters in the game version need longer to be processed and result in lower sensitivity. Gamification of the ambiguous word task also reduced performance. One explanation is that this difference is due to the expectations participants bring with them into the experiment; within the magic circle, games can been experienced as a place where aggression is expected and players are given license to act in aggressive ways, potentially increasing their access to hostile words when in a game environment.

For player experience, we see no differences in the cognitive tasks; however, we find differences in the social and aggression tasks that point to more motivating and positive experiences in the task compared to the game. One explanation is that prior expectations of game scenarios influence players of the game – players expect a game to fulfill certain standards based on prior experience. For gamified tasks in general, the game mechanics and aesthetics may not match expectations, diminishing the experience. For experimental tasks, this is not true,

Figure 5 The game version of the executive functioning task

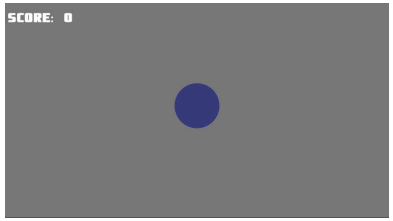

Figure 6 The task version of the executive functioning task

Figure 7 The game version of the facial recognition task

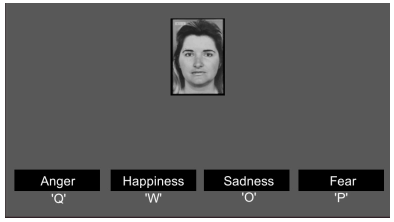

Figure 8 The task version of the facial recognition task

because we don't expect a task to stimulate fantasy or give interesting problems. In our case, the two affected conditions still felt like experimental tasks and less like a game (compared to the cognitive games), perhaps raising expectations for gameplay that were not met.

Conclusion

Opportunities: The promising take-away message from our results is that changes in experience don't necessarily affect performance. This opens up several design opportunities to create psychological tasks that can benefit from game-elements without threatening the reliability of the task. Psychological tasks can be used to assess different populations, e.g., children or older adults, by providing appeal. Because of the higher entertainment value, players may be more motivated to start and to complete multiple repetitions of a gamified task. This is often necessary, when for example, cognitive functions are being monitored.

Challenges: It might be more challenging than previously assumed to integrate game design into an already well-designed task. Considering that the mechanic is at the core of each game, gamifying the task through the addition of premise might not add as much value to simple tasks as we assumed. Additional research is necessary to understand how we can best integrate game-elements into psychological tasks to maximize benefits and minimize reliability issues.

Strategies: To create strategies that don't violate the reliability of the task, it is important to understand that adding game-elements might not foster a more positive experience, depending on the context and purpose of a task. Designers should ensure that the task is evaluated and compared to the gold standard.

Ethics: Not ensuring that a task is comparable to the gold standard has ethical implications in both research and therapeutic contexts because people are often classified based on their task performance. Additionally, embedding performance measures into a game-context has ethical implications, because games can be used for assessment without explicit participant consent.

References

[1] Anderson, C.A., Carnagey, N.L., & Eubanks, J. (2003). Exposure to violent media: The effects of songs with violent lyrics on aggressive thoughts and feelings. JPSP, 84, 960-971.

[2] Cohn, J. F., Schmidt, K., Gross, R., & Ekman, P. (2002). Individual differences in facial expression: Stability over time, relation to self-reported emotion, and ability to inform person identification. In *Proc. of MI* (p. 491).

[3] Geurts, L., Vanden Abeele, V., Van Keer, K., & Isenborghs, R. (201). Playfully learning visual perspective taking skills with sifteo cubes. In *Proc. CHI play'14* (pp. 107-113).

[4] Kane, M. J., & Engle, R. W. (2002). The role of prefrontal cortex in working-memory capacity, executive attention, and general fluid intelligence: An individual-differences perspective. *Psychonomic bulletin & review*, *9*(4), 637-671.

[5] Khatib, F., DiMaio, F., Cooper, S.... & Foldit Void Crushers Group. (2011). Crystal structure of a monomeric retroviral protease solved by protein folding game players. *Nature structural & molecular biology*, *18*(10), 1175-1177.

[6] Mandryk, R. L., Dielschneider, S., Kalyn, M. R.... & Keiver, K. (2013). Games as neurofeedback training for children with FASD. In *Proc. IDC'13*, 165-172.

[7] Nicholson, S. (Forthcoming). A RECIPE for Meaningful Gamification. To be published in Wood, L & Reiners, T., eds. Gamification in Education and Business, New York: Springer

[8] Simmonds, D. J., Pekar, J. J., & Mostofsky, S. H. (2008). Meta-analysis of Go/No-go tasks demonstrating that fMRI activation associated with response inhibition is task-dependent. *Neuropsychologia*, *46*(1), 224-232.

[9] Von Ahn, L., & Dabbish, L. (2008). Designing games with a purpose. *Communications of the ACM*, *51*(8), 58-67.