Using Game-Based Environments to Measure Cognitive Decision Making

Laura A. Waters, Karen L. Blackmore

Using Game-Based Environments to Measure Cognitive Decision Making

Laura A. Waters[1 [0000−0002−6432−2338]] and Karen L. Blackmore[1 [0000−0002−9111−0293]]

The University of Newcastle, New South Wales 2308, Australia
laura.a.waters@uon.edu.au

Abstract. Within the area of serious games research, there is significant potential for researchers and other stakeholders to use serious games to gain more fundamental understanding of the underlying cognitive processes of individual users or participants. In this research, we present the results of an experiment to benchmark a visual search task presented in a 3d game-like environment with a standard, controlled, lab based implementation. Our results show similar trends in performance measures across experimental conditions in the two environments, however, participants were faster and more accurate overall in the 3d game-like environment. There is significant potential for researchers and other stakeholders to utilise serious games platforms as a means of measuring human cognition within environments that are visually more closely related to 'real-life' than those used in cognitive psychology.

Keywords: 3D environment · cognition · visual search task · decision making.

1 Introduction

Serious games (SGs), or games that are used for a purpose other than purely entertainment, are being used more frequently within industries for assessment and training of specific skills and abilities. While some performance measures may be captured, there is great benefit in understanding the types of cognitive processes that underlie abilities. One important research area of cognitive psychology is decision-making; that is, how we process information to make a judgment. This is often done through lab-based experiments that ask participants to make simple cognitive judgments about simple perceptual stimuli.

Many SG models acknowledge that at the crux of any learning activity, there is a cognitive process occurring [7, 15]. However, the incorporation of robust measures of cognitive performance in game based environments is not currently done, even though the captured data might shine a light on understanding the processes underlying their cognitive abilities [3, 7, 11]. Within the field of computer science, there has been extensive research into 2D environments and evidence has shown that visual search abilities improve with gaming [2], however these findings were often incidental outcomes and visual search abilities have not been explicitly considered in an immersive game-based environment in a way the relates to the more formal, constrained approach used in cognitive psychology [6].

In this work we present the results of a pilot study to compare visual search task performance, as measured through reaction time (RT), between a standard 2D constrained lab based task, and a 3D game-like visual environment. The focus here is to assess whether the deviations in RTs that occur when the number of visual stimuli change maps between the 2D lab based task and the 3D counterpart. We begin by providing a brief background to the measurement of human cognition, followed by an outline of our research focus, the experimental methodology and the results of our pilot study. Lastly, we provide some concluding discussion, and outline potential avenues for future work.

2 Background

2.1 Measuring Human Cognition

Cognitive psychology uses robust methodological approaches to test an individuals underlying processing abilities and capacity when performing particular cognitive tasks. There are a number of robust tools and tasks that are used to understand cognitive performance that can provide this broader understanding of capability; for example, surveys, psychometrics, observation, randomised controlled trials (RCTs). One key difficulty in being able to generalise task performance to real world behaviours lies in the highly constrained nature of lab-based environments and artificial stimuli used [5], such as simple shapes. These concerns relate to the overall context of the results, and the potential impact of this context on performance [8]. On the other hand, studies of entertainment games and SGs tend to use quasi-experimental designs and surveys and largely tend to measure post-intervention activity engagement [12]. While studies using quasi-experimental designs and surveys have added to our understanding of the outcomes and impacts of playing games, RCTs provide more rigorous evidence about the impacts of games. Although, laboratory-based tests provide a robust foundation for our understanding of decision making processes, more research is required for conversion of highly constrained lab based tasks and measures into other contexts such as more realistic virtual games-based environments. The rise of SGs and immersive simulation environments provides a platform for a more realistic and variable environment which could be used for cognitive testing and training purposes. Simplistically, performance data can be used to profile players or measure competency [1], and there is promise for game environments to be used to capture more cognitive level data to assess cognitive processing architecture and capacity.

2.2 Research Focus

At the higher level, the purpose of this research is to develop the foundations of a theoretical framework that uses the strengths of two currently separate research areas to address some of their respective limitations. On one hand, cognitive psychology uses robust methodological approaches to design, measure performance, empirically analyse, and make inferences about the underlying cognitive processes and abilities that people use to complete a task or make decisions. We understand that human cognition plays a role in nearly every aspect of every day life, however, due to the highly constrained nature of lab-based tasks, it can be difficult to generalise results to how an individual might approach tasks in their everyday environment. On the other hand, the gaming industry has conducted extensive research into human-computer interaction, engagement and player typing/preferences. By combining aspects of SGs with cognitive lab-based approaches, we aim to develop a framework which provides guidelines to both researchers and game designers on a) how to increase engagement and realism when developing experiments in cognitive psychology, and b) how to embed robust cognitive methodology in a SG. This would allow for finer grain data and a greater depth of understanding about how individuals are processing and engaging within the environment. This can help us to adapt or optimally design games for learning to accommodate cognitive styles. This is particularly relevant for SGs, where these games are typically deployed across a cohort of individuals with different player profiles and styles.

One overall end-goal of this research will consider how virtual environments impact users; that is, do people behave and make decisions in a more realistic and virtual environment in the same way they would in standard artificial lab-based environments. However, in order to retain the robustness and validity of any cognitive measures used in a game-based environment, we first need to lay the groundwork for ensuring reliable measures of in-game performance by benchmarking performance on simple tasks using the highly constrained approach of cognitive psychology. Although we are restricting participant interaction with the 3D game-like environment, the goal of this pilot study is to ensure that the more visually immersive environment is still capturing valid and reliable data that can be used as a performance baseline for future studies in which we allow greater interactive possibilities for participants. As a means of benchmarking performance on a cognitive task between a standard 2D lab-based cognitive psychology test and a 3D game-like environment, we elect to design and pilot a simple visual search task and replicate this in a restricted game-like environment. Python programming language was used to create the 2D lab-based task whereas we used Virtual Battlespace 3 (VBS3) [4], a visually realistic game environment with scenario creation tools, to create the 3D game-like environment.

We hypothesise that participants will have a preference for the mechanics used in the immersive environment of VBS3, and that this will be reflected in this pilot experiment through faster RTs and greater accuracy even though the cognitive processes involved should remain the same. If we find evidence to support the hypothesis that game-based environments are a valid way to deliver cognitive tasks and assess underlying cognitive processes, then we are developing the capacity to re-deploy this back into the design of SGs and potentially allow for greater flexibility in the realism of environments used for cognitive testing.

3 Methodology

In this pilot study, we designed a simple visual search task [14] within a 3D virtual environment (Virtual Battlespace 3) which replicated the experimental design of a typical 2D laboratory-based task requiring participants to make a simple decision. Further details on the environments and visual stimuli are provided in the materials section below. The research was conducted under the University of Newcastle's Human Research Ethics Committee approval number H-2018-0227.

3.1 Participants

Participants were undergraduate students from the University of Newcastle. Of the 41 students who completed the experiment, there were 28 males (M = 22.23 years, SD = 3.71) and 13 females (M = 25.38 years, SD = 9.91) ranging from 18 to 52 years of age. Participants were enrolled in either a second year software engineering course (28 males and 2 females respectively) or a first year introductory psychology course (11 females) and received course credit for their participation.

3.2 Materials and Design

The experiment was conducted in one of the University of Newcastle HCI labs using an Alienware 17 R3 laptop. PsychoPy2 (v1.90.3) was used to develop and run the 2D laboratory-based task, while the 3D game-based task was run using VBS 3.7.0.

Design In each of the environments, participants were presented with either 5, 10, 15 or 20 items (stimuli) in a single 4 second trial and required to make a simple decision on whether a target item was present or not. The desired target depended on the presented colour arrangement of an item. Stimuli used in standard cognitive experiments are often highly constrained shapes or letters which are varied visually through the use of colours, location, or rotation [14]. Replication of our task in a 3D game-like environment meant that to achieve a higher level of realism, abstract visual stimuli could not be used in both conditions. As such, this experiment used human avatars in VBS3 to address the criteria of perceptual salience between targets and distractor items, relying on colour arrangement of clothing to provide distinction. In each environment, the target (2D: a rectangular, two toned shape, and 3D: a human avatar, see Figure 1 and 2 for example trials) would be tan on top and red on the bottom, while both distractor types would be tan on the bottom half, and either red or yellow on top.

For each task, a practice block with feedback (correct or incorrect response) was completed prior to undertaking 10 experimental blocks, each containing 40 trials (no feedback). We counterbalanced task and response key orders to account for any potential ordering effects. We also asked participants to complete two (2) questionnaires; one pre-test demographic survey containing additional questions regarding gaming preferences and behaviours, and one post-test where we asked participants about the difficulty of the two tasks and whether they were aware of using a particular search strategy to reach their decision.

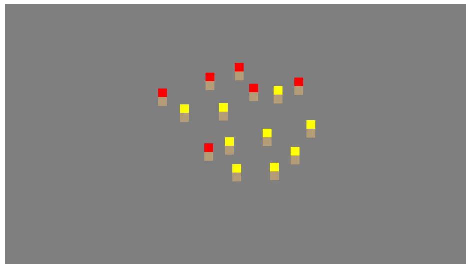

Fig. 1. An example 15-item set presented in the 2D environment. The target item is absent from this trial.

Fig. 2. An example 20-item set from the VBS3 (3D) environment. The target avatar is present on this trial.

4 Results

4.1 Preliminary Analysis

Surface level analysis of this pilot experiment revealed a significantly faster mean RT (M = 0.74s) in the 3D environment compared to the 2D environment (M = 1.02s), t = 74.515, $p < .001$, as well as higher accuracy overall (M = 99% and M = 97% for 3D and 2D respectively), t = -9.191, $p < .001$. At first glance it could be proposed that in each task participants respond 'yes' as soon as the target item has been identified (self-terminating cognitive architecture) on target present trials, or exhaustively search all items in target absent trials before responding 'no'. [13]. This is reflected in both 2D and 3D environments by the slower average RTs on trials where no target item was present (see below figure).

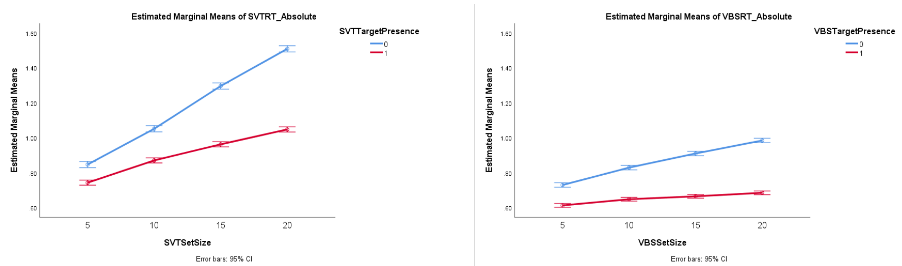

Fig. 3. The absolute RTs averaged across participants can be shown in the two graphs above (Left graph: RTs for the 2D environment, Right graph: RTs for the 3D VBS3 environment). Participants were significantly faster when completing the task in VBS3 compared to the 2D python environment, however in both tasks were participants tended to respond slower to trials without a target present.

5 Discussion

Two main arguments in favour of using game-like or scenario based platforms for teaching and training purposes are that firstly, they are more engaging and therefore more effective as a training tool [9], and secondly, they can be more closely related to the real world meaning an individuals performance is more likely to reflect actual behaviour [10]. Unfortunately, the reported benefits and effectiveness of these training approaches are often mostly subjective rather than objectively measured, and assumptions are often made about generalisability to real world behaviours.

The findings of this simple pilot experiment indicate that embedding cognitive tasks within virtual training environments with high levels of 'realism' hold promise as a comparative environment for the capture of robust measures of underlying cognitive processes. The group level trends of slower RTs on trials where target item was absent or when there was a greater number of items to be processed (as seen in figure 3), indicates that participants were engaging similar underlying cognitive decision-making processes despite the different environments. While the design of the task itself was highly constrained, it is worth noting that the faster response times and improved accuracy in the VBS3 environment may be impacted by additional factors such as a discrepancy in the perceived perceptual planes of the two environments. This factor warrants further investigation and as such, future experimentation will focus on adjusting the locations of the avatars as well as the viewing angle of the 'player avatar' so the spatial locations are more closely aligned with the 2D plane presented in the comparative environment. By further researching performance measures within these 2D and 3D comparative environments, we aim to develop a foundation for designing and implementing tasks within SGs that can provide rich, robust and valid measures of a users cognitive processing abilities. This would also provide researchers, educators, or game designers with greater insight into the cognitive abilities, interactions and learning styles of specific SG users.

References

[1] Marc Busch et al. "Using Player Type Models for Personalized Game Design-An Empirical Investigation." In: *Interaction, Design and Architecture* 28 (2016), pp. 145–163.

[2] Alan D Castel, Jay Pratt, and Emily Drummond. "The effects of action video game experience on the time course of inhibition of return and the efficiency of visual search". In: *Acta psychologica* 119.2 (2005), pp. 217–230.

[3] Glenda Gunter, Robert F Kenny, and Erik Henry Vick. "A case for a formal design paradigm for serious games". In: *The Journal of the International Digital Media and Arts Association* 3.1 (2006), pp. 93–105.

[4] Bohemia Interactive. *Virtual Battlespace 3*. 2015. URL: `https://bisimulations.com/products/virtual-battlespace`. Accessed: 2019.

[5] Simon Ladouce et al. "Understanding minds in real-world environments: toward a mobile cognition approach". In: *Frontiers in human neuroscience* 10 (2017), p. 694.

[6] Moon-Soo Lee et al. "Characteristics of Internet Use in Relation to Game Genre in Korean Adolescents". In: *CyberPsychology and Behavior* 10.2 (2007). PMID: 17474846, pp. 278–285. DOI: `10.1089/cpb.2006.9958`.

[7] Igor Mayer et al. "The research and evaluation of serious games: Toward a comprehensive methodology". In: *BJET* 45 (2014), pp. 502–527.

[8] Geoff Norman. "Generalization and the qualitative–quantitative debate". In: *Advances in Health Sciences Education* 22.5 (Dec. 2017), pp. 1051–1055. ISSN: 1573-1677. DOI: `10.1007/s10459-017-9799-5`.

[9] Clark N Quinn. *Engaging learning: Designing e-learning simulation games*. John Wiley & Sons, 2005.

[10] Ute Ritterfeld, Michael Cody, and Peter Vorderer. *Serious games: Mechanisms and effects*. Routledge, 2009.

[11] Valerie J. Shute. *Learning Processes and Learning Outcomes*. English. Distributed by ERIC Clearinghouse [Washington, D.C.], 1992, 38 p. URL: `https://eric.ed.gov/?id=ED366660`.

[12] Shamus P. Smith, Karen Blackmore, and Keith Nesbitt. "A Meta-Analysis of Data Collection in Serious Games Research". In: *Serious Games Analytics: Methodologies for Performance Measurement, Assessment, and Improvement*. Ed. by Christian Sebastian Loh, Yanyan Sheng, and Dirk Ifenthaler. Cham: Springer International Publishing, 2015, pp. 31–55. ISBN: 978-3-319-05834-4. DOI: `10.1007/978-3-319-05834-4_2`. URL: `https://doi.org/10.1007/978-3-319-05834-4_2`.

[13] James T. Townsend. "Serial vs. Parallel Processing: Sometimes They Look like Tweedledum and Tweedledee but they can (and Should) be Distinguished". In: *Psychological Science* 1.1 (1990), pp. 46–54. DOI: `10.1111/j.1467-9280.1990.tb00067.x`.

[14] Jeremy M Wolfe. "Visual attention". In: *Seeing*. Elsevier, 2000, pp. 335–386.

[15] A. Yusoff et al. "A Conceptual Framework for Serious Games". In: *2009 Ninth IEEE International Conference on Advanced Learning Technologies.* July 2009, pp. 21–23. DOI: 10.1109/ICALT.2009.19.