Integration of a Cognitive Assessment Task into Exergame Gameplay Elements

Mana Carr
University of Auckland
mcar397@aucklanduni.ac.nz

Anthony Dang
University of Auckland
adan849@aucklanduni.ac.nz

Alex Shaw
School of Computer Science
University of Auckland
l.shaw@auckland.ac.nz

Burkhard Wuensche
School of Computer Science
University of Auckland
burkhard@cs.auckland.ac.nz

Abstract

Lack of exercise is related to a variety of health issues, including cognitive decline in older adults. Tools for encouraging regular exercise such as exergames are a useful preventative measure, but regular screening for impairment is still important. However, standard cognitive screening methods can be both time-consuming and tedious. Integrating these screening methods into a frequently played exergame is one way to enable regular screening, but requires that the integration is not obtrusive and does not interfere with the gameplay.

We present an exergame in which a standard cognitive screening tool, the AX-Continuous Performance Task, is integrated into the game in a non-obtrusive fashion. As an starting step in assessing this approach, we validate the comparability of the measurement capacity of this integrated tool by assessing user performance in the test with non-impaired adults. Our results indicate that the test is comparable to the traditional form of the test when conducted within the context of a game, and is not clearly perceived as a test rather than a gameplay element by the users. However, increasing task complexity through additional gameplay elements does interfere with task performance.

1. Introduction

Exercise has a significant health benefit for people of all ages. Conversely, a lack of exercise is associated with a range of health problems, particularly as people age [1, 2]. However, adherence to exercise is often low, even in the case of medically prescribed exercise programs [3]. One possible solution to aid in counteracting this problem is with the use of exergaming, which is a combination of exercise and video gaming [4]. By combining the two together, the motivation to exercise may be increased by introducing enjoyable game-play elements.

One health issue often faced by people as they age is cognitive decline [5]. As people grow older, cognitive ability begins to deteriorate, and this is associated with reductions in both quality and length of life [6]. Detecting when and if cognitive ability starts to deteriorate is an ongoing issue within the medical field, and the quicker cognitive decline can be detected, the more effective mitigation strategies can be. The detection of mental aptitude and decline can be achieved with the use of cognitive testing and screening.

Regular screening for cognitive decline is important for early detection. However, standard cognitive tests are generally tedious and time-consuming, meaning that adherence to any self-managed screening program is likely to be low. In a similar manner to how exergames can increase exercise adherence, the integration of standard cognitive tests into video games could be a useful tool for increasing screening frequency. Some mechanics from existing video games such as the mobile game Fruit Ninja already appear to be effective tests of cognitive functions [7].

We propose that exergames offer high potential in this area: as games with the intention that users are playing them regularly for fixed amounts of time in order to promote physical fitness, they are a good avenue for the integration of cognitive screening. This then means that by playing, the user will receive both regular screening, and also the preventative benefits of exercise to mitigate any potential decline.

Virtual Reality (VR) is an area that has seen some recent exploration in the field of exergames, and shown some potential in distracting the user from their exercise. Immersive VR delivered via a head-mounted-display also offers potential for cognitive screening: by completely controlling the user's experienced environment, interference from external factors can be minimised and thus the integrity of the test results improved.

Our research objective is to determine whether tools for cognitive screening can be integrated into an exergame in a non-obtrusive manner while remaining

an effective test. In particular, we aim to validate task performance in a gameplay-integrated test against a non-integrated one to determine whether the integrated test will give comparable indicators of executive function.

2. Background

There is a wealth of research on the effects of exercise on cognitive functions [8, 9, 10]. In this paper, we instead focus specifically on exergaming and VR. There has been relatively little literature examining either exergaming or modern, immersive VR as venues for cognitive testing and screening.

2.1. Exergaming With Cognitive Testing

There is a lot of existing material on the effects of exercise on cognitive ability, particularly with the use of exergaming. This work has primarily focused on integrating cognitive training into exercise to boost or preserve cognitive capabilities and quality of life, rather than for the purpose of screening for decline.

In one example, past work investigated whether exergaming in short bursts of time had any noticeable cognitive effect on the individuals in the study [11]. They conducted the study by implementing both an engaging exergame with the cognitive test, and a second version of the cognitive game which involved no physical activity. It was found that people that took the test with the exergame had an acute cognitive performance boost over those that did not play the exergame. It also found that individuals exerted the same amount of energy in both test cases but considered the exergame version to be more fun.

Anderson et al. sought to determine whether a cycling-based exergaming would improve cognitive function, prevent cognitive decline, and increase participation in more physical activities [12]. The research paper determined that older adults that participated in the cycling exergame had achieved better cognitive ability than those who had cycled normally, while exerting the same amount of effort. It also concluded that the results of this experiment suggest that there is some merit with the use of simultaneous cognitive and physical exercise to prevent cognitive decline.

The research of Karssemeijer et al. aimed to investigate the effects that a 12-week long program, involving exergame training and aerobic training separately, have on a controlled group of people suffering from dementia induced frailty [13]. The paper concluded that the 12-week exergame training program produced positive effects on the control group, by reducing the level of frailty in people with dementia. Also, it is noted that the individuals involved in the exergaming program were more inclined to do more physical exercise, which suggests that exergaming is an effective method to get people engaged with more physical exercise.

Lastly, a paper reviewed the use of computer-assisted screening to determine whether it would aid neuro-psychological assessment [14]. The authors found that the use of computers in this screening reduced the amount of time taken to complete the assessment. It also increased the accuracy of such assessments and made cognitive screening a more readily available tool for professionals. This suggests that the use of exergaming for cognitive screening has a lot of potential to be an effective tool for screening.

2.2. Virtual Reality Based Cognitive Testing

There has been a substantial amount of work that examines the use of non-immersive virtual environments in healthcare [15]. This past work has demonstrated that virtual environments are suitable for a variety of purposes, though the research tends to focus more on education and interventions, rather than testing and assessment. There is however comparatively little work on the use of immersive VR in healthcare, and very little in the area of cognitive testing and screening. Much of what does exist also predates the rise of modern, highly immersive heads-up-displays, such as the Oculus Rift.

The use of VR with cognitive testing has been increasing in usage [16]. This meta-analysis concludes that VR can be used to determine the cognitive health of a participant. It achieved this conclusion by studying the many other existing studies covering the use of VR for cognitive testing, focusing on the sensitivity of the measures taken. The work of Rose et al. also gives a similar conclusion on the use of VR on brain damage rehabilitation [17]. It even states that the use of VR in brain damage rehabilitation is expanding dramatically and may become an integral part of cognitive testing within the near future.

Chua et al. investigated the use of VR as a cognitive screening tool, whether it was feasible and accepted by the individuals involved in its testing [18]. They also investigated whether the VR screening module itself was able to distinguish cognitively impaired people from cognitively intact people. The individuals involved in this test were mainly older people coming from a primary care setting. The authors concluded that there was a positive correlation between the results from the VR screening module and the traditional cognitive screening method, and that the module was generally

accepted by those who used it. This research suggests that implementing a cognitive test within an exergame is feasible.

Other research also investigates the effectiveness and feasibility of VR, but regarding testing spatial navigation memory in older individuals [19]. It compared the use of an immersive VR-based environment with a similar screening environment displayed on a desktop computer. Notable outcomes from this research project were that the participants in the VR-based environment were more engaged with the screening, made less navigational mistakes, and were able to recall more environmental landmarks than those that participated in the computer-based setup. Like the previous research paper, it also concluded that VR is a feasible tool to use with cognitive screening.

3.   Design

3.1.   Original Exergame Design

In this research, we integrate the AX-CPT into the mechanics of an existing VR exergame previously shown to be effective at eliciting high intensity exercise [20]. The original exergame contained several game elements such as cannons, sand pits, beams and power-ups, as demonstrated in Figure 1. This exergame makes use of a computer-connected exercycle as the exercise component of the gameplay.
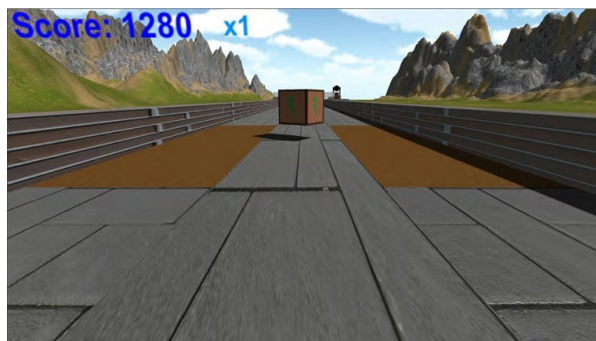

Figure 1.   Base exergame with gameplay elements visible.

The game environment is automatically generated as the player cycles through the map. The score is kept on the top left, alongside a multiplier that affects any points awarded. The player must move along the track, avoiding the obstacles (cannons, overhead beams, sandpits) by moving their body, and picking up power-ups.

3.2.   AX-Continuous Performance Task

We chose the AX-Continuous Performance Task (AX-CPT) as the cognitive test to integrate into the game. This test is a good choice as it measures aspects of cognition known to show decline and change with age, and is sufficiently complex to make for non-trivial game mechanics.

The AX-CPT is a cognitive test that measures an individuals sustained cognitive control. It is a test that measures an individual's ability to exercise both proactive and reactive cognitive control. Proactive control is the ability of an individual to anticipate an event and react appropriately despite primed interference. Reactive control is the ability to detect interference after is has occurred and react appropriately despite it. Successful cognition depends on being able to exercise both. Older adults typically see a shift towards reactive control as the ability to exercise proactive control diminishes with age [21, 22]

The events come to test subjects in the form of alternating cue and probe lettering pairs. Subjects are first shown a cue letter, and must respond appropriately, then a probe letter, and again must respond appropriately. Participant reactions are either a positive or negative response, usually implemented as pressing either of two buttons. Cue letters are categorised as A or B, while probe letters are categorised as X or Y. Usually for A and X, there is only one letter option, while B and Y instances may be B, Y or any letter that is not A or X. This sequence is shown in Figure 2.

In the case of the cue, subjects must always give a negative response. In the case of the probe, subjects respond depending on both the current probe letter, and the cue letter that immediately preceded it. If the cue-probe combination matches the target sequence (A-X), the user gives the positive response. If it does not match, either due to the probe not being an X, or the preceding cue not being an A, the user gives the negative response.

The sequence A-X makes up the majority of cue-probe pairs, and as such subjects are naturally primed into an alternating negative-positive response sequence. When a pair occurs that does not match the sequence, the subject must exercise cognitive control to provide the correct response. B-X pairs test proactive control, as the subject must use the fact that they had previously seen a B cue in order to not give the positive response upon seeing the X. A-Y pairs test reactive control, as the subject must react appropriately upon seeing the Y probe. Subjects have only a 1-1.5 second window in which to give a response, forcing them to

exhibit control rather than spend time thinking about the appropriate response.

Figure 2. AX-Continuous Performance Task Cue-Probe Sequence

3.3. AX-Continuous Performance Task Integration

We build upon previous research that implements the AX-CPT within the chosen exergame [23]. In this previous work, the AX-CPT is not integrated with the game mechanics and was found to be quite obtrusive when playing, as visible in Figure 3.
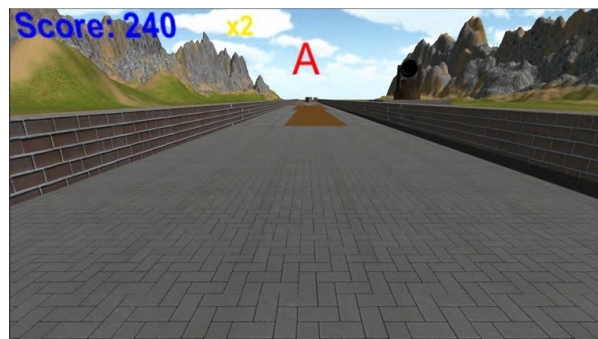

Figure 3. AX-CPT prompts visible in exergame.

To integrate the AX-CPT into the game, we replace the obtrusive AX-CPT display with a flying drone that follows the player and fire upon them. Since the drone fires upon the player, the cannons were made obsolete and were removed from the exergame.

The drone implements the cue-probe cycle of the AX-CPT through alternating a charging phase and a firing phase. During the charging phase, the drone absorbs energy, corresponding to A or B based on the colour. During the firing phase, it absorbs some more energy before firing either a laser or power-up at the player. While in the charging phase, the player should give a response with the left trigger to ready themselves and charge up for the firing phase. The drone will either absorb green or blue energy in this instance. Once in the firing phase, the player should activate their shield if the drone starts emitting red energy after absorbing green energy in the charging phase. Otherwise, the player should use the left trigger to use their power-up magnet to retrieve a power-up. The magnet and shield are illustrated in Figure 4.
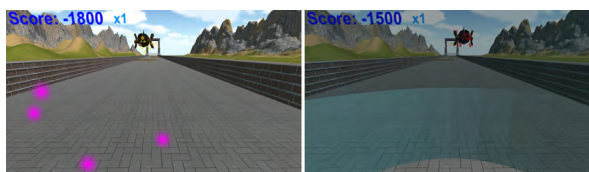

Figure 4. Responses from player perspective. Magnet response is left, shield response is right.

The colours on the drone are analogous to the cue and probe prompts in the original AX-CPT test, with green representing "A" and red representing "X". The target sequence in this implementation is therefore green-red. Blue and yellow colours represents a letter that was not of the target sequence.

One key difference that this implementation has over the traditional AX-CPT test is that the player is notified if they have given the correct response. This is untrue within the original implementation of the AX-CPT test, as the player would not know whether they gave the correct response or not. This is a necessary and unavoidable change in order to make a meaningful gameplay element. If there was no indication that responding to the drone did anything, there would be no motivation to respond to the drone firing on the player.

4. Implementation

This work was implemented using Unity version 2018.2.8f1.

An Oculus Rift CV1 headset was used as the Virtual Reality (VR) equipment that participants would need to use to enter the VR Environment. Sensors were placed around the exercycle to for a balanced run-through of the exergame. The Oculus Rift controllers were used for providing positive or negative responses to test stimuli. The controllers were attached to the bike behind the front heart-rate sensors to allow participants to reach the triggers while also holding onto the handlebars. This was important to miminise the risk of a participant falling off the bike or losing balance as they navigate through the exergame. This setup is demonstrated in Figure 5.

The game also makes use of a depth-camera-based head and body tracking system to enable the user to steer through body motion and duck under obstacles.

5. Evaluation

A cross-sectional within-subjects study was conducted to evaluate the integration of the AX-CPT into the exergame gameplay.

The study contained three test conditions, which

Figure 5. Oculus Rift controllers affixed for use on exercycle.

participants completed in a counterbalanced order. All three conditions required participants to cycle on the exercycle during the test period.

5.1. Base Condition

The first test condition, the Base Condition, consisted of the original AX-CPT implementation inside an exergame environment that contained no game elements. Figure 6 shows an illustration of the described game environment.
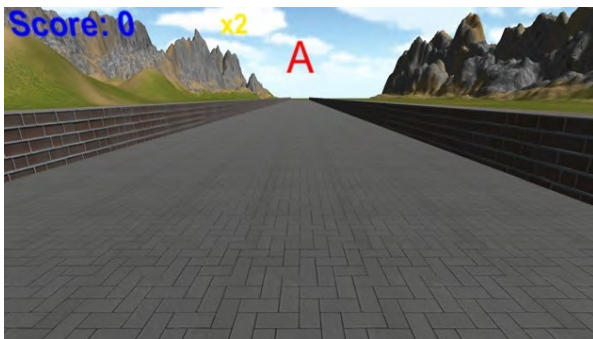

Figure 6. The base test condition, containing no gameplay elements and showing an "A" cue from the test.

5.2. Gameplay Condition

The second condition, the Gameplay Condition, consisted of the same original AX-CPT test as the first condition, but within an exergame environment which had all the game elements previously mentioned. A screenshot of this condition is shown in Figure 7.

Figure 7. The gameplay test condition, containing gameplay elements and showing an "X" probe from the test.

5.3. Integrated Condition

The last test condition, the Integrated Condition, was our implementation of the AX-CPT test using the drone as a substitute for the letter prompts, within the same exergame environment as the Gameplay Condition, including gameplay elements. This is the implementation of the discreet AX-CPT test to be compared to the previous two test conditions. The drone and it's behaviour is illustrated in Figure 8.

Figure 8. The integrated test condition, containing a drone which substitutes for gameplay elements and implements the AX-CPT. The red particles indicate the drone is charging an attack, which is the equivalent of an "A" cue.

Each test condition contained two blocks of 50 cue-probe pairs each, with each block separated by a 30 second break in order to reduce focus degradation. In total, each condition had 100 pairs.

5.4. User Study Procedure

Participants were recruited through an open call to university students. Of the respondents, 23 met the inclusion/exclusion criteria of lacking any health

issues which could cause a problem for the exercise component of the test, and being able to use an immersive VR head-mounted-display without suffering from cybersickness. Of those 23 people, 14 identified as male, and 9 identified as female. The age of participants ranged between 20 and 25 years, with a median age of 22. Participants came from a variety of ethnic backgrounds.

Participants self-reported a mean 3.48 hours of moderate-to-high intensity exercise per week, and a mean 3.57 hours of video game play time per week. On a scale from 1-10, participants reported their levels of physical and mental exhaustion prior to beginning the study as means of 4.88 and 5.47 respectively.

Young adults were chosen for the study to minimise the chances of undiagnosed dementia disorders or other aspects of age-related cognitive decline being present in the participant cohort. While the end purpose of integration of cognitive testing into gameplay is to target older adults, it is important to establish a baseline equivalence between the integrated test and the traditional test prior to measuring the effect in older adults or those suffering from dementia disorders.

Informed consent was obtained from all participants, and the study was approved by the DELETED FOR REVIEW Ethics Committee (reference number: DELETED FOR REVIEW).

As a validation of the comparability of task performance rather than an intervention study, this work uses a single-session, within-subjects design. This is in order to directly compare participant performance across conditions.

Participants completed a pre-test questionnaire to provide general demographic data such as age and sex, and self-report measures including time spent in exercise, time spent playing video games, perceived level of fitness, and current perceived level of physical and mental exhaustion.

Following the questionnaire, participants were given an instruction sheet explaining the exergame and each of the three test conditions. After reading this, they completed a short tutorial on the AX-CPT, delivered outside of VR on a desktop computer with the E-Prime test software, and given time to get comfortable with the test. This was to minimise learning effects with respect to test performance.

After completing the AX-CPT tutorial, the participants also completed a trial run of the exergame, with no test present. This allowed them to get comfortable with the VR environment and with the basic in-game movements, such as steering and ducking under beams. This time was also used to calibrate the headset, head-tracking tools, and exercycle seat position.

The participants then completed the three test conditions in a counterbalanced order, determined by the method of Latin Squares. Prior to beginning each condition, participants were reminded of the nature of the condition in question. Each test condition had a duration of 8 minutes, and participants had a mandatory break with a minimum duration of 5 minutes before beginning the next condition.

Finally, after completing all test conditions, participants were asked to rank each condition in terms of how mentally exhausting it was, on a scale from 1-10. Answering this question marked the end of the session.

5.5. Measures

We focus on two main measures related to AX-CPT performance:

Error Rate: The number of times a participant responds incorrectly to either a cue or a probe. The four main errors are derived from the division of cue-probe pairs: AX, AY, BX, and BY. While we also record whether these errors occur on the cue or the probe response, only the probe response gives information directly related to cognitive control. Incorrect responses to BX probes correspond to failures of proactive cognitive control, while incorrect responses to AY probes correspond to failures of reactive cognitive control.

Response Rate: The frequency at which participants respond or fail to respond to either cues or probes within the 1 second window in which they are visible.

6. Results

Data analysis was achieved using an Analysis of Variance (ANOVA) test. The ANOVA was used to compare the data between each of the test conditions and determine if there are any significant differences amongst the test conditions error rates.

6.1. Overall Performance Across Conditions

Two sets of comparison data were produced during our analysis. One set of data shows the comparison between the overall performance in each of the three conditions, in terms of error rates. This comparison shown in Figure 9. The data in this data set only accounts for cue-probe pairs where participants did provide a timely response. The lower the error rate, the fewer errors that were discovered in that condition.

A second set of data was also collected which also compared the overall performance of each condition against each other. However, in this data set, cases where participants did not provide a timely response to the probe prompt were also included in the analysis. The results from this comparison are shown in Figure 10.

As we can see from the data presented in Figure 9, the Integrated condition has the worst performance overall in terms of error fractions. There is a negligible difference between the Gameplay and Integrated conditions, and thus with this data set, it can be determined that the AX-CPT implementation in the Integrated Condition is somewhat comparable to the implementation in the Gameplay condition. Although both conditions are comparable to each other, they both are significantly different to the Base condition. It can be determined that they both are not comparable to the implementation in the Base condition.

Furthermore, when cases with no response to probe prompts are considered in the data analysis, similar results are achieved. All three test conditions have a higher error ratio in this data set, than the data set shown when only explicit incorrect responses were analysed. It is interesting to note that the Gameplay and Integrated conditions are now more comparable to each other, as the differences between both conditions has decreased. Although the Base condition's error rate has increased, both of the other conditions have increased by a greater amount, becoming less comparable to it.
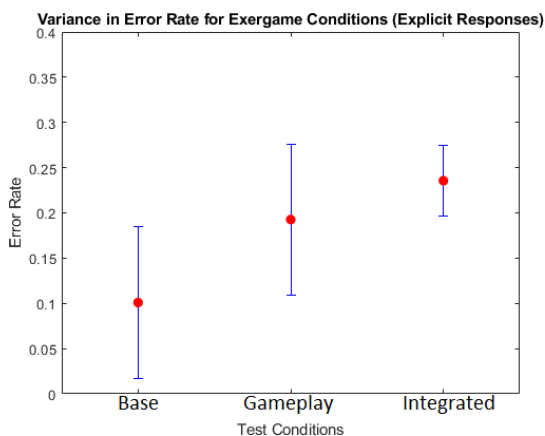

Figure 9. Error rates for each condition, considering explicit probe responses only.

6.2. Performance Across Error Types

Along with the overall performance comparison, a comparison on the performance of the participants for each of the cue-probe pair types was also conducted. The data that was collected for this comparison was
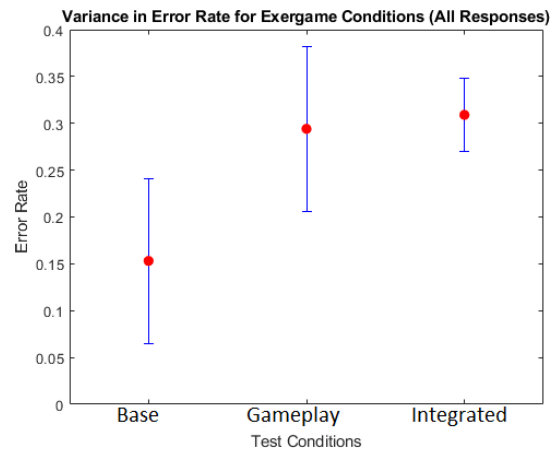

Figure 10. Rates of either incorrect or missing probe responses for each condition.

independent of test condition. This is again divided both into a dataset of only explicit incorrect responses, and one of both incorrect and missing responses. The first set is shown on the graph in Figure 11.
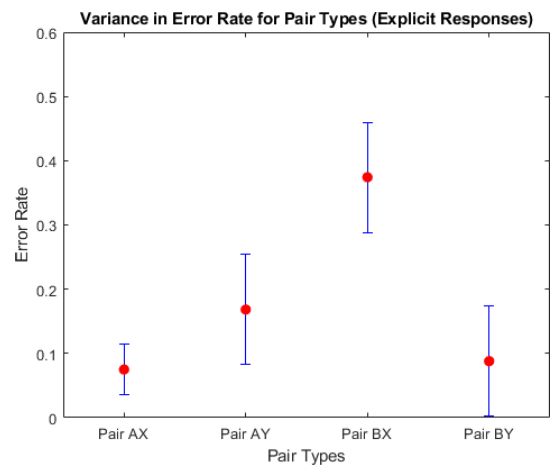

Figure 11. Error rates for explicit probe responses, divided by cue-probe pair types.

Figure 12 shows the results with both incorrect and missing responses. Like the analysis on overall performance, once all responses are taken into account, error rates for each compared type increase but retain the same general pattern.

Both the data sets show that the BX pair type is the worst performing pair type of all. What this tells us about the participants in this study is that they demonstrated poor proactive control when encountering the pair type in question. It is most likely that the participants did not proactively register that they are not seeing the target sequence when they encounter the B cue. Once they have seen the X probe, they give an

incorrect response since they assume that they have just seen the target sequence.
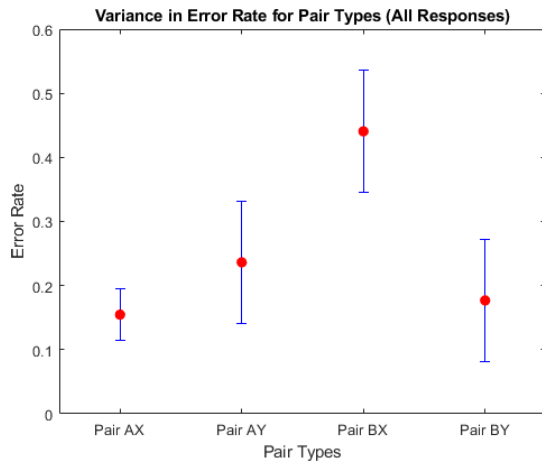

Figure 12. Rates of incorrect or missing probe responses, divided by cue-probe pair types.

Test Case	Mean Mental Exhaustion
Base Condition	3.57
Gameplay Condition	6.30
Integrated Condition	5.57

6.2.1. Perceived Mental Exhaustion Table 1 shows the mean ratings for mental exhaustion given to each of the conditions. The Gameplay Condition contained the highest scores in terms of mental exhaustion, with the Integrated Condition following. This meant that the Gameplay Condition was the most tiring out of the three conditions. The Integrated Condition was also determined to be somewhat mentally exhausting, but not to the same degree as the Gameplay Condition. The Base Condition was ranked the lowest in terms of mental exhaustion amongst the study participants. The differences are statistically significant (F: 18.60, $p < 0.01$). It is interesting to see that although both the Gameplay and Integrated conditions are somewhat comparable in terms of performance, the Integrated Condition was considered less mentally exhausting than the Gameplay one. Comments made by the participants during the end of the user study commonly described the Base Condition as the most uninteresting of the three tests. Participants also mentioned that for the non-Integrated conditions, they did not see the point of responding to the cue-probe prompts, as nothing was shown to indicate that they gave a correct or incorrect response.

7. Discussion

Our results indicate that both the Gameplay and Integrated conditions require significantly more cognitive effort than the Base condition. This is consistent with previous findings [23]. As expected, the requirement of completing two mental tasks simultaneously significantly increases the difficulty.

The Integrated condition was not described as a test to participants. This fact may have contributed to the performance dip between the Base and Integrated conditions. The implementation of the AX-CPT test was obvious to the player in the Base and Gameplay conditions. Participants understood that they were being subjected to a cognitive test. This was not the case for the Integrated condition as it was described as a game, and the cognitive test itself was described in terms of game elements instead of its traditional test elements. This may have impacted performance as participants may not have felt the added pressure of undergoing a test, and therefore put in less effort into responding to the drone correctly. In the conditions with an unambiguous AX-CPT, the participants may have placed higher importance on the test elements than on any other components of the virtual environment (particularly in the Gameplay condition), while in the Integrated condition the drone may have been assigned importance comparable to other elements in the environment.

Next, another contributing factor that could have affected the results is the fact that the AX-CPT was implemented to give explicit feedback in the Integrated condition. Since the participant is given feedback on whether they gave the correct response to the drone, they may be more inclined to give the correct responses. If they give an incorrect response, they know immediately, and this may cause people to give another incorrect response. The other conditions do not give explicit feedback, and therefore participants do not realize they had given the incorrect response and will continue with the test uninterrupted. Participants do not see any consequences if they give an incorrect responses in those conditions.

One interesting note is that some participants commented did not initially recognise the game implementation of the AX-CPT as being the same test as the other implementations, even when performing the game version shortly after having had the test explained to them, or after completing two other versions of the test. This is a good indicator that the integration of the test is not obtrusive to the user.

The fact that many participants did not recognise

the game as a test also has promise for applications of cognitive testing outside of dementia screening. In particular, if someone knows they are being subjected to a cognitive test, their performance is likely to differ to that shown in their regular activities. Thus, for measuring functional cognitive performance, having the ability to test it without the subject being consciously aware of the test could be very useful.

7.1. Limitations

There are several limitations affecting our findings. One such threat is the fact that in the Integrated condition, the AX-CPT test needed to be altered slightly to provide immediate feedback to the player, as traditional gameplay elements depend on feedback. This was a necessary change in order to give players a reason to interact with the drone. This feedback is not present in the other two conditions, nor the traditional AX-CPT test.

The degree to which the cues and probes were interpretable may have been impacted by their implementation. Colours were used as substitutes of the traditional lettering in the Integrated condition. This may have been a factor to the performance dip seen in the Integrated condition.

There were some usability issues associated with the use of the Oculus Rift headset. In particular, an update by Oculus that mandated the use of the external sensors for tracking caused occasional conflicts with the head and body tracking system used by the exergame when the two reported differing movements from the participants. This could manifest as a slightly misaligned view in the game, causing some discomfort.

Our results may not generalise to all populations. While it is necessary to first test with cognitively healthy individuals in order to validate the game integration, it is possible that dementia disorders may affect game-playing capabilities in such a way that the test results are not comparable between conditions.

8. Conclusion

To conclude, this research has found that it is feasible and effective to integrate a cognitive test discreetly into an exergaming environment. Although the discreet implementation of the AX-CPT test had significantly lower performance than the implementation with no game elements, it does however possess comparable performance results with a tradional implementation that does contain said game elements. Along with that result, participants in the user study also rated the discreet implementation as less mentally exhausting than the implementation with game elements. This

suggests that the AX-CPT test integrated discreetly as game elements is likely to be played more and be a more effective screening mechanism.

However, due to the difficulty of completing both gameplay tasks and testing tasks, our results indicate that tests should be integrated into games such that they make up the main, or only, gameplay elements, so to avoid other elements interfering with the test.

8.1. Future Work

There still remains much work to pursue in this line of research. We see two primary directions in which this research should be extended:

Firstly, the adaptation of this work for seniors and those suffering from age-related cognitive decline or dementia disorders. This work has demonstrated that discreet integration of cognitive testing into gameplay is viable and produces useful results in a young, healthy population. While it is likely that these results extend to older adults an additional study targeting such adults is required to confirm this. This research would also benefit from further investigation from a game design perspective: how can a test-integrated game be made most suitable for older adults, especially those whose capacities may be reduced. Finally, similar work from an exercise persepctive would also greatly benefit this line of work, as an exercycle-based exergame is only appropriate for subjects with a certain degree of physical capability.

The second main avenue of further research we see is the broadening of this work to other cognitive tests - both for assessing cognitive control in other ways for further validation of results, and also for assessing other cognitive functions. The N-back test could be a good choice, both due to its relatively varied possible implementations and widespread use in traditional cognitive assessment. It would be interesting to discover whether there are suitable adaptable tests for all executive functions or whether certain ones cannot be appropriately integrated.

References

[1] F. W. Booth, C. K. Roberts, and M. J. Laye, "Lack of exercise is a major cause of chronic diseases," *Comprehensive physiology*, vol. 2, no. 2, pp. 1143–1211, 2011.

[2] N. Owen, G. N. Healy, C. E. Matthews, and D. W. Dunstan, "Too much sitting: the population-health science of sedentary behavior," *Exercise and sport sciences reviews*, vol. 38, no. 3, p. 105, 2010.

[3] F. Jones, P. Harris, H. Waller, and A. Coggins, "Adherence to an exercise prescription scheme: the role of expectations, self-efficacy, stage of change and

psychological well-being," *British journal of health psychology*, vol. 10, no. 3, pp. 359–378, 2005.

[4] D. E. Warburton, S. S. Bredin, L. T. Horita, D. Zbogar, J. M. Scott, B. T. Esch, and R. E. Rhodes, "The health benefits of interactive video game exercise," *Applied Physiology, Nutrition, and Metabolism*, vol. 32, no. 4, pp. 655–663, 2007.

[5] T. Hedden and J. D. Gabrieli, "Insights into the ageing mind: a view from cognitive neuroscience," *Nature reviews neuroscience*, vol. 5, no. 2, pp. 87–96, 2004.

[6] J. K. Johnson, L.-Y. Lui, and K. Yaffe, "Executive function, more than global cognition, predicts functional decline and mortality in elderly women," *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, vol. 62, no. 10, pp. 1134–1141, 2007.

[7] Q. Liu, X. Zhu, A. Ziegler, and J. Shi, "The effects of inhibitory control training for preschoolers on reasoning ability and neural activity," *Scientific reports*, vol. 5, no. 1, pp. 1–11, 2015.

[8] K. Davranche and T. McMorris, "Specific effects of acute moderate exercise on cognitive control," *Brain and cognition*, vol. 69, no. 3, pp. 565–570, 2009.

[9] T. McMorris, J. Sproule, A. Turner, and B. J. Hale, "Acute, intermediate intensity exercise, and speed and accuracy in working memory tasks: a meta-analytical comparison of effects," *Physiology & behavior*, vol. 102, no. 3-4, pp. 421–428, 2011.

[10] C. H. Hillman, M. B. Pontifex, L. B. Raine, D. M. Castelli, E. E. Hall, and A. F. Kramer, "The effect of acute treadmill walking on cognitive control and academic achievement in preadolescent children," *Neuroscience*, vol. 159, no. 3, pp. 1044–1054, 2009.

[11] Y. Gao and R. Mandryk, "The acute cognitive benefits of casual exergame play," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 1863–1872, 2012.

[12] C. Anderson-Hanley, P. J. Arciero, A. M. Brickman, J. P. Nimon, N. Okuma, S. C. Westen, M. E. Merz, B. D. Pence, J. A. Woods, A. F. Kramer, *et al.*, "Exergaming and older adult cognition: a cluster randomized clinical trial," *American journal of preventive medicine*, vol. 42, no. 2, pp. 109–119, 2012.

[13] E. G. Karssemeijer, W. J. Bossers, J. A. Aaronson, L. M. Sanders, R. P. Kessels, and M. G. O. Rikkert, "Exergaming as a physical exercise strategy reduces frailty in people with dementia: a randomized controlled trial," *Journal of the American Medical Directors Association*, vol. 20, no. 12, pp. 1502–1508, 2019.

[14] R. F. White, K. E. James, J. J. Vasterling, R. Letz, K. Marans, R. Delaney, M. Krengel, F. Rose, and H. C. Kraemer, "Neuropsychological screening for cognitive impairment using computer-assisted tasks," *Assessment*, vol. 10, no. 1, pp. 86–101, 2003.

[15] R. Ghanbarzadeh, A. H. Ghapanchi, M. Blumenstein, and A. Talaei-Khoei, "A decade of research on the use of three-dimensional virtual worlds in health care: a systematic literature review," *Journal of medical Internet research*, vol. 16, no. 2, p. e47, 2014.

[16] A. Neguț, S.-A. Matu, F. A. Sava, and D. David, "Virtual reality measures in neuropsychological assessment: a meta-analytic review," *The Clinical Neuropsychologist*, vol. 30, no. 2, pp. 165–184, 2016.

[17] F. D. Rose, B. M. Brooks, and A. A. Rizzo, "Virtual reality in brain damage rehabilitation," *Cyberpsychology & behavior*, vol. 8, no. 3, pp. 241–262, 2005.

[18] S. I. L. Chua, N. C. Tan, W. T. Wong, J. C. Allen Jr, J. H. M. Quah, R. Malhotra, and T. Østbye, "Virtual reality for screening of cognitive function in older persons: comparative study," *Journal of medical Internet research*, vol. 21, no. 8, p. e14821, 2019.

[19] K. Ijaz, N. Ahmadpour, S. L. Naismith, and R. A. Calvo, "An immersive virtual reality platform for assessing spatial navigation memory in predementia screening: Feasibility and usability study," *JMIR mental health*, vol. 6, no. 9, p. e13887, 2019.

[20] L. A. Shaw, B. C. Wünsche, C. Lutteroth, S. Marks, J. Buckley, and P. Corballis, "Development and evaluation of an exercycle game using immersive technologies," 2015.

[21] T. S. Braver, J. L. Paxton, H. S. Locke, and D. M. Barch, "Flexible neural mechanisms of cognitive control within human prefrontal cortex," *Proceedings of the National Academy of Sciences*, vol. 106, no. 18, pp. 7351–7356, 2009.

[22] K. Jimura and T. S. Braver, "Age-related shifts in brain activity dynamics during task switching," *Cerebral Cortex*, vol. 20, no. 6, pp. 1420–1431, 2010.

[23] A. Shaw, J. Buckley, P. Corballis, C. Lutteroth, and B. Wuensche, "Exploring the possibility of virtual reality exergaming as a cognitive screening system," in *Proceedings of the 53rd Hawaii International Conference on System Sciences*, 2020.