Multi-player online video games for cognitive rehabilitation

Rationale for experimental design

Researcher: Jason Colman
Supervisors: Dr Jim Briggs, Dr Louise Turner, Dr Alice Good, Dr Paul Gnanayutham
Version 1.1

August 2013
# Table of Contents

1 Introduction ................................................................................................................................. 3  
   1.1 Background and rationale for the hypothesis ........................................................................ 3
2 Research design summary .......................................................................................................... 3
   2.1 Previous studies ................................................................................................................. 4
3 Rationale for general research design decisions ....................................................................... 4
   3.1 Single study research design (SSRD), not a randomised controlled trial (RCT) ............... 4
   3.2 Computer-based cognitive tests ....................................................................................... 4
   3.3 Cognitive tests are integrated into the game software ...................................................... 4
   3.4 Length of study ................................................................................................................. 5
   3.5 Session duration ................................................................................................................ 5
   3.6 Phase duration .................................................................................................................... 5
   3.7 Invitation-only game .......................................................................................................... 5
   3.8 Why develop a game - why not use an off-the-shelf game? ............................................. 5
4 The cognitive tests ....................................................................................................................... 6
   4.1 What are the tests measuring? ......................................................................................... 6
   4.2 Validity of computerised versions of the tests .................................................................. 6
   4.3 Test instructions ............................................................................................................... 6
   4.4 Test practice modes ......................................................................................................... 6
5 Reaction time tests ..................................................................................................................... 6
   5.1 Computerised reaction time tests ..................................................................................... 6
   5.2 Implementation in present study ..................................................................................... 7
6 Stroop tests .................................................................................................................................. 7
   6.1 Computerised Stroop tests .............................................................................................. 7
   6.2 Stroop Word ....................................................................................................................... 8
   6.3 Stroop Colour ..................................................................................................................... 8
   6.4 Colour-Word ...................................................................................................................... 8
7 Trail-making tests ....................................................................................................................... 8
   7.1 Computerised trail-making tests ....................................................................................... 9
   7.2 Sequential numeric trail-making test (task “A”) ............................................................... 9
   7.3 Alternating number and letter trail-making test (task “B”) ............................................... 9
8 Rationale for game design decisions ......................................................................................... 9
   8.1 Game design summary ...................................................................................................... 9
   8.2 Why 2D? ............................................................................................................................ 9
   8.3 Why the low resolution graphics? ..................................................................................... 9
   8.4 Why do the players have customisable avatars? ............................................................... 9
   8.5 Why are the avatars humanoid? ....................................................................................... 9
9 References ..................................................................................................................................... 9
1 Introduction
The aim of this research project is to find out if multi-player online video games can provide a form of cognitive therapy, for brain-injured people. The study focuses on people who have survived a traumatic brain injury (TBI) or stroke.

The hypothesis is that *playing a multi-player online video game could provide a form of cognitive therapy for a brain injured person.*

1.1 Background and rationale for the hypothesis
Brain injuries afflict an estimated 10 million people per year worldwide (Hyder et al., 2007), a problem described by Theodoros et al. as having reached “epidemic proportions” (2001, p. 27). According to Mateer (2005), successful rehabilitation tends to be comprised of multiple eclectic approaches, again structured around the needs of the individual. The aim of this study is to find a new, complimentary path to regaining lost skills.

According to Sohlberg & Mateer, a basic assumption underlying cognitive rehabilitation is that cognition cannot be treated in isolation: “Brain damage affects cognitive, social, behavioural and emotional functioning” (2001, p. 10). Ben-Yishay & Daniels-Zide (2000) suggest that acceptance of the disability leads to a life which is more emotionally satisfying than one in which comparisons with pre-injury abilities are constantly being made. This is interpreted by Sohlberg & Mateer (2001) as implying that cognitive and emotional recovery for a brain injured person are inseparable. Cicerone et al. (2000) also note that cognitive rehabilitation may incorporate interventions aimed at improving a person's emotional functioning.

Jones et al. (2010) assessed 630 individuals with an acquired brain injury, and found a surprising positive relationship between injury severity and life satisfaction. The authors show that the strengthening of personal identity and social relationships are beneficial for ABI survivors, and conclude “...individuals can be protected from the negative impact of more severe head injury by receiving support from social networks and by strengthening personal identity.”

Perceived social isolation - loneliness - has a dramatic effect on social animals, including man (Cacioppo & Hawkley, 2009). Loneliness is a risk factor for cognitive decline, and isolation is common following brain injury (Murdoch & Theodoros, 2001). Playing an MMO provides a social experience, even for those who play alone (Ducheneaut et al., 2006). McGonigal (2011) states 'social network games make it easier and more fun to maintain strong, active connections with people we care about'.

Massively multi-player online games (MMOs) are a combination of video game and virtual environment where social interaction with other players is the norm. The findings of Jones et al. (2010) that the social network of a brain-injured person contributes to his or her quality of life, lead to the conclusion that playing an MMO could be a form of cognitive therapy, by fostering collaboration, cooperation, and social contact.

2 Research design summary
This section summaries the research design. The Experimental Protocol document gives full details.

The study will employ a single-subject research design (SSRD) for each participant. An SSRD relies on repeated measurements of the dependent variable (Christensen, 2004). The experiment follows an “ABAB” design, according to the schedule shown in table one: participants alternate
weeks of playing an online multi-player video game with weeks where the game is not played. Every Friday, participants take a series of cognitive tests.

<table>
<thead>
<tr>
<th>Monday</th>
<th>Tuesday</th>
<th>Wednesday</th>
<th>Thursday</th>
<th>Friday</th>
<th>Saturday</th>
<th>Sunday</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Play game</td>
<td>Play game</td>
<td>Play game</td>
<td>Play game</td>
<td>Take tests</td>
<td>Play game</td>
<td>Play game</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Play game</td>
<td>Play game</td>
<td>Play game</td>
<td>Play game</td>
<td>Take tests</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1. Schedule for one participant.

2.1 Previous studies
Malec et al. (1984) performed a study with a small group of brain-injured people, to determine if playing a video game provided a form of cognitive therapy. The game was single-player only (Target Fun on the Atari VCS). Cognitive improvement was measured using Stroop tests, letter cancellation tests, and a reaction time test. A statistically significant improvement in reaction time was found; the Stroop test and letter cancellation test results did not show a significant improvement.

The design of the current study has been based to some extent on the study of Malec et al. (1984). This will allow a comparison with that study (evaluating a single-player game) and the present study (evaluating an online multi-player game).

3 Rationale for general research design decisions
This section describes the rationale for the experimental design.

3.1 Single study research design (SSRD), not a randomised controlled trial (RCT)
Because the number of participants is expected to be low, and because of the very variable nature of brain injury, a randomised controlled trial design is not considered appropriate. Instead, a single-subject research design (SSRD, or “n=1” design) is to be employed (Janosky, 2005). In this design, time series data is collected for each participant as they experience different interventions.

3.2 Computer-based cognitive tests
In this study, the cognitive tests taken by the participants are administered by the game software, with the intention that the physical presence of the researcher is not required. The reason for this is to increase the number of potential participants, in two ways: by removing physical location as a barrier, and by removing the need for a person to administer the test to each participant.

3.3 Cognitive tests are integrated into the game software
The cognitive tests are not only computerised, but are also part of the game software, rather than being one or more stand-alone applications. This is for the following reasons:
- To simplify the task of downloading and running the software for the participants: there is only one piece of software.
- The integration between tests and game allows the software to require the participant to take the tests before the game can be played further. This may incentivise the participant to take the tests, when they may otherwise be skipped – i.e. this may improve compliance.
- The tests may be made more fun by giving the participant rewards for taking the tests (or
improving in them), which can be used in the game.

3.4 Length of study
This study employs an “ABAB” design, where each phase (“A” or “B”) takes one week. The entire study takes four weeks and one day, for each participant. Participants do not all have to start on the same day, and so the entire study may be run for longer, with individual participants joining the study after it has started.

For participants who cannot commit to the 4-week schedule, a shorter, 3-week schedule could be used (“ABA” design). The reason for allowing this shorter version is to maximise the number of participants who will be able to commit to the study.

3.5 Session duration
Lerdal et al. (2009) state that fatigue is a common complaint following a stroke, but knowledge regarding post stroke fatigue is limited.

Session length is to be kept as short as possible, to minimise fatigue. This must be balanced with the need to keep the sessions long enough for any potential effects to develop, and for the participants to complete the weekly cognitive tests. Therefore the first run of the experiment will be a pilot to determine a reasonable session length. In an unstructured interview with one participant, it was found that 15 minutes may be the maximum acceptable duration, but this is expected to be widely variable. Session lengths are recorded on the server, which may be used to guide the next run of the experiment.

3.6 Phase duration
One phase in this design is a single week, during which the participant either plays the game for four days, then takes the cognitive tests on the fifth; or does not play the game that week, only taking the tests on the Friday.

This scheme follows that used by Malec et al. (1984).

3.7 Invitation-only game
Gaggioli et al. (2007) caution that using virtual environments for online therapy exposes participants to risks arising from their anonymous and open nature. Privacy, impersonation of others, and unwanted intrusion are all issues which must be considered, leading these authors to suggest the use of private, dedicated servers to host the virtual environment.

These findings were corroborated in unstructured interviews between potential participants and the researcher. Therefore, only invited participants may log in to the game and interact with other players.

3.8 Why develop a game - why not use an off-the-shelf game?
Developing a game allows for all aspects of the game experience to be customised to the requirements of the experimental design and the special needs of the participants. Off-the-shelf commercial games are certainly entertaining, but are not research instruments. The following advantages arise from being able to change any part of the game software:

• The game can be made as simple as possible, so it is not overwhelming for the participants to play.
• Allows us to have single- and multi-player versions of the game. If two completely different games were used for these two experimental phases, there would be much more scope for confounds.
• Allows accessibility features to be added to the software as required by the participants.
Commercial game software tends to not be very accessible.

- Allows the server-side software to be customised as required. For instance, every change in player position is logged, so that the movements of a player can be replayed and analysed.

4 The cognitive tests

In this study, it is proposed to use the following cognitive tests as measures.

- Reaction time
- Stroop tests
- Trail-making tests.

The dependent variables in the study are the results of these tests. The independent variable is whether or not the participant has played the online multi-player game in that week.

Malec et al. (1984) used letter and symbol cancellation tests. These tests have been discarded from the present study as they are used to diagnose spatial neglect, rather than being a general measure of executive function etc. (Rorden & Karnath, 2010).

4.1 What are the tests measuring?

In each case, it is proposed that the test measures a cognitive ability that may potentially be influenced by playing the online multi-player game.

4.2 Validity of computerised versions of the tests

In this study, cognitive tests are automatically administered by the combined game/test software, rather than manually by a physically present researcher. For each test, rationale for the validity of the computerised version of the test is given.

In general, the validity of computer-based cognitive tests has been shown in commercial and non-commercial offerings.

4.3 Test instructions

Short instructions are given before each test, every time it is run.

4.4 Test practice modes

The tests as implemented in the present study have a practice mode, where the participant may have a trial run of the test. An arrow points to the (or one) correct choice, demonstrating how to successfully complete the test.

The following sections give the rationale for the inclusion and implementation details of each of the cognitive tests.

5 Reaction time tests

Playing video games invariably improves the reaction time of the player (Griffiths, 2005), including the brain-injured participants in the study of Malec et al. (1984). A reaction time test is included in the present study primarily as a control: it is expected that an improvement in reaction time will be seen, before any improvement in other tests. If no improvement in reaction time is seen, this may indicate that a parameter of the experiment needs adjustment – e.g. the session duration has not been long enough to cause any improvement.
5.1 Implementation in present study

Malec et al. (1984) describe their reaction time test as follows.

*Reaction time was time between a light stimulus and a button press response in milliseconds. A spoken “ready” signal was given prior to stimulus presentation.*

In the present study, the user is presented with the word “ready” on screen (this is spoken by the computer if text-to-speech is enabled). After a random delay of a few seconds the word “go” appears, accompanied by a sound effect. After this point the time taken for the participant to click on the screen is measured.

6 Stroop tests

The Stroop colour-word test is a commonly used measure of executive function. It measures the ability to shift cognitive set, and is believed to measure cognitive inhibition (Homack & Riccio, 2004) - the ability to suppress a learned response in favour of an unusual response.

Homack & Riccio point out that there are numerous variants of the Stroop tests and no “official” version. In the present study, three separate tests are given: the Word, Colour, and Colour-Word tests. This follows Malec et al. (1984).

6.1 Computerised Stroop tests

The validity of using computerised Stroop tests is supported by the findings of Hepp et al. (1996), who compared a computerised Stroop test with a manually administered version, in which the test was printed on cards, and responses given verbally. The computer microphone was used to measure reaction time, but responses were recorded manually by a researcher. These authors found that participants sometimes blended words together, (e.g. “gred”, or “bleen”), meaning that some responses could not be counted as correct or incorrect. This suggests that a computerised/pointing version could be more accurate than oral responses. Overall, Hepp et al. recommend that “computerized versions of the Stroop task should be used.”

Examples of computer-based Stroop tests can be found online, e.g. [http://www.snre.umich.edu/eplab/demos/st0/stroop_program/strooptestshockwave.dcr](http://www.snre.umich.edu/eplab/demos/st0/stroop_program/strooptestshockwave.dcr); (requires ShockWave plug-in.) This particular test gives the participant two choices for each test, so there is a 50-50 chance of guessing correctly.

Stroop tests are implemented in PEBL. In the PEBL implementation, participants have four options for each test, and respond by pressing one of the keys 1-4.

In the present study, four user interface “buttons” are presented on screen (figure 1). The button order does not change, following the PEBL Stroop test design, where the key mapping does not change. For each test, the correct button is (likely to be) different to the previous time. On a computer with a mouse, the participant must move the mouse cursor to the chosen button, then physically press the mouse button. On a touch screen device, the participant must move his or her finger to their choice and touch the screen. Thus there is a motor response time as well as a cognitive response time which contributes to the overall time taken to make a choice. (This is also the case for other computerised tests which do not react to a verbal response). Van der Elst et al. (2006) found that speed-dependent scores on Stroop tests are 'profoundly' affected by demographics. It is therefore more accurate to use the proportion of errors, rather than the time to complete the test.
Figure 1. Computerised Stroop test in the present study

6.2 Stroop Word test
Malec et al (1984) describe this test as follows.

*Stroop word requires the subject to read columns of randomly sequenced color words (“red”, “blue”, “green”) as rapidly as possible. The score is number of words read in 45 seconds.*

In the present study, the number of correct and incorrect choices are recorded. To avoid the issues noted above of separating motor reaction time from cognitive response time, the ratio of correct to incorrect scores will be used in analysis of the results.

6.3 Stroop Colour test
Malec et al (1984) describe this test as follows.

*Stroop Color requires naming columns of print in three randomly sequenced colours (red, blue, green).*

6.4 Stroop Colour-Word test
Malec et al (1984) describe this test as follows.

*The Color-Word task requires naming the color of the ink of discrepant color words (e.g. saying “red” to describe the word “blue” printed in red ink). The score here is total correct in 45 seconds.*

7 Trail-making tests
Trail-making tasks are widely used to measure cognitive flexibility and attention, as well as visual capabilities (Zakzanis et al., 2005). In the manual version of these tests, the participant uses a pencil to connect randomly distributed numbered or lettered circles in the stipulated order. In the simple case, the circles contain numbers and must be connected in sequential numeric order. In the second task, the circles contain alternating numbers and letters. The circles should be joined in the order 1-A-2-B-3-C, etc. These are commonly called tasks “A” and “B”.

7.1 Implementation of trail-making tests
Figure 2 shows the implementation of this test in the present study.
Instead of moving a pencil on paper, the participant moves the mouse cursor, or their finger in the case of a touch screen device. Motor coordination and speed is still a factor in the results. The computerised test immediately alerts the participant to an incorrect choice, whereas the response time from a human examiner may vary (Zakzanis et al., 2005).

**8 Rationale for game design decisions**

In this section, rationale is presented for the design of the game developed as the research instrument in this study.

**8.1 Game design summary**

Players inhabit a shared environment containing food, treasure, and harmful enemies which deplete players' health. Health is restored by eating food, but players can only eat food given to them by other players. Thus players are reliant on each other to eat and stay alive. Players can communicate by typing text into a shared chat window. This game design is intended to foster communication, collaboration, and cooperation. It is a 'socially significant' situation as described by Ray (2004, p. 56), intended to heighten the emotional involvement for players.

**8.2 Why 2D?**

Conversations with potential participants revealed that some brain-injured people find 3D scenes difficult to process, or feel sea-sick viewing them.

**8.3 Why the low resolution graphics?**

It is the hypothesis of the researcher that group communication and community-based aspects of an online multi-player game that would provide any cognitive improvements, rather than the graphical aspect of the game. This study is based to some degree on the study of Malec et al. (1984) which used the Atari VCS, primitive by modern standards. The low resolution look of the present software is intended to be 'retro' enough to allow a comparison – using photorealistic graphics would be a confound. (As an aside, this style is also contemporary in some games, e.g. Minecraft is very low-resolution.)

**8.4 Why do the players have customisable avatars?**

Some participants wanted their avatars to look like them, or to look like some other well known character. So customisable avatars were accepted as a software requirement.

**8.5 Why are the avatars humanoid?**

Play testing and a questionnaire revealed that the overwhelming majority of players wanted human(oid) avatars, as opposed to non-humanoid characters such as animals.
9 References


