Multi-player online video games for cognitive rehabilitation for the brain injured

Jason Edward Colman

This thesis is submitted in partial fulfilment of the requirements for the award of the degree of Doctor of Philosophy of the University of Portsmouth.

September 2016
Table of Contents

List of tables ........................................................................................................................................... v
List of figures ............................................................................................................................................ vi
Declaration ................................................................................................................................................ viii
Abbreviations .......................................................................................................................................... ix
Abstract ................................................................................................................................................... x
Dissemination ........................................................................................................................................... xi
Publications .............................................................................................................................................. xi
Presentations ........................................................................................................................................... xii
Conferences ............................................................................................................................................. xii
Acknowledgements ............................................................................................................................... xiii

1 Introduction ........................................................................................................................................... 1
  1.1 Motivations ....................................................................................................................................... 2
  1.2 Research questions and hypotheses ................................................................................................. 3
  1.3 Research approach ........................................................................................................................... 4
  1.4 Aims and objectives ......................................................................................................................... 5
  1.5 Original contribution to knowledge ................................................................................................. 5
  1.6 Structure of the remainder of this dissertation ................................................................................ 5

2 Background and research questions .................................................................................................. 7
  2.1 Acquired Brain Injury (ABI) ........................................................................................................... 7
    2.1.1 Stroke ....................................................................................................................................... 7
    2.1.2 Traumatic Brain Injury (TBI) .................................................................................................. 9
    2.1.3 Effects of brain injury .............................................................................................................. 10
    2.1.4 Assessment of brain injury ...................................................................................................... 11
  2.2 Rehabilitation and quality of life after brain injury ......................................................................... 12
    2.2.1 Cognitive rehabilitation .......................................................................................................... 12
    2.2.2 Cognitive remapping (neuroplasticity) .................................................................................... 13
    2.2.3 Quality of life ......................................................................................................................... 13
    2.2.4 The disability paradox ........................................................................................................... 15
    2.2.5 Loneliness .............................................................................................................................. 15
    2.2.6 Summary ................................................................................................................................... 16
  2.3 Therapeutic uses of video games .................................................................................................... 17
    2.3.1 Flow, play and games .............................................................................................................. 17
    2.3.2 Video games ............................................................................................................................ 18
    2.3.3 Video games for physical and behavioural therapy ............................................................... 19
    2.3.4 Video games for cognitive therapy .......................................................................................... 20
    2.3.5 Multi-player and massively multi-player online games ....................................................... 20
    2.3.6 Comparing virtual environments, virtual reality and MMOs ................................................ 21
    2.3.7 Therapeutic uses of MMOs and virtual environments ....................................................... 22
    2.3.8 Summary ................................................................................................................................... 23
  2.4 Conclusions and questions ............................................................................................................ 23

3 Towards a Research Design ................................................................................................................ 26
  3.1 Research methodologies ................................................................................................................... 26
    3.1.1 Research designs .................................................................................................................... 26
    3.1.2 Sample size .............................................................................................................................. 26
    3.1.3 Participant recruitment ........................................................................................................... 27
    3.1.4 Single-subject research design, SSRD ................................................................................... 28
    3.1.5 Statistical analysis of SSRDs ................................................................................................ 29
    3.1.6 Proposed experimental design ............................................................................................... 29
  3.2 Cognitive tests ................................................................................................................................... 30
    3.2.1 The Mini-Mental State Examination (MMSE) ....................................................................... 31
    3.2.2 Clock-drawing test (CDT) ...................................................................................................... 34
Experiment #1: Measuring cognitive improvement

1. The participants
   1.1 Recruitment activities
   1.2 Participant details

2. Method
   2.1 Schedule

3. Results
   3.1 Stroop tests
   3.2 Trail-making tests
   3.3 Reaction time test
   3.4 ‘Percentage change towards normal’ analysis
   3.5 Log in data

4. Discussion

5. Conclusions

Experiment #2: Measuring wellbeing

1. Changes to the experimental protocol
2. Changes to the game design
   2.1 New ‘community’ screens
   2.2 New trading mechanic

3. The ‘daily digest’ player-incentivising emails

4. The new measures
   4.1 De Jong Gierveld Loneliness Scale
   4.2 Satisfaction With Life Scale

5. Method
   5.1 Participant recruitment
   5.2 Schedule

6. Results
   6.1 Engagement and compliance
   6.2 Loneliness Scale
   6.3 Satisfaction With Life Scale (SWLS)
   6.4 Stroop tests
   6.5 Trail-making tests
   6.6 Reaction time
   6.7 ‘Percentage change towards normal’ analysis
   6.8 Qualitative results

7. Discussion
   7.1 Engagement
   7.2 Loneliness
   7.3 Satisfaction With Life
   7.4 The cognitive tests
   7.5 Practice effects and the Hawthorne effect
   7.6 Follow-up test results
   7.7 Qualitative responses

8. Summing up

Conclusion

1. Conceptual model
2. Research approach
3. Game software design and development
4. Experiments
List of tables

Table 2.1. Definition of stroke (Sacco et al., 2013) .............................................................. 8
Table 3.1. Scoring manual for the clock-drawing test on 0-5 scale (Kørner, Lauritzen, Nilsson, Lolk, and Christensen, 2012) ................................................................. 35
Table 3.2. De Jong Gierveld 6-point Loneliness Scale questions ........................................ 41
Table 3.3. Satisfaction With Life Scale questions ................................................................. 42
Table 6.1. Game design features, both included and discarded, and their intended purpose. 83
Table 7.1. Recruitment activities for 2013 study ................................................................. 93
Table 7.2. Brain-injured participants ...................................................................................... 94
Table 7.3. Schedule for 2013 experiment ............................................................................. 97
Table 7.4. Test session dates ............................................................................................... 99
Table 7.5. Stroop word results .......................................................................................... 100
Table 7.6. Stroop colour results ......................................................................................... 100
Table 7.7. Stroop colour-word results .............................................................................. 100
Table 7.8. Trail-making (A) results .................................................................................... 102
Table 7.9. Trail-making (B) results .................................................................................... 102
Table 7.10. Reaction time results ...................................................................................... 104
Table 7.11. Value of ‘X’ (‘lower limit of normal’) for use in ‘percentage change towards normal’ analysis ................................................................. 106
Table 8.1. Schedule for experiment #2 ............................................................................ 125
Table 8.2. Dates of phases in experiment #2 ................................................................... 126
Table 8.3. Number of active participants in each phase .................................................. 127
Table 8.4. Number of test sessions per participant ......................................................... 128
Table 8.5. Total loneliness for each participant, for each phase. Brain-injured participants are marked ‘ABI’ .......................................................................................... 131
Table 8.6. Social Loneliness scores .................................................................................. 132
Table 8.7. Emotional Loneliness scores ......................................................................... 132
Table 8.8. Mean Satisfaction With Life scores for each participant, for each phase .......... 135
Table 8.9. Stroop word results ........................................................................................ 138
Table 8.10. Stroop colour test results .............................................................................. 139
Table 8.11. Stroop colour-word test results .................................................................... 139
Table 8.12. Trail-making (A) test results ........................................................................ 143
Table 8.13. Trail-making (B) test results ........................................................................ 143
Table 8.14. Reaction time test results ............................................................................ 146
Table 8.15. Value of ‘X’ for use in ‘percentage change towards normal’ analysis .......... 149
Table 8.16. Qualitative responses to the question ‘Did you feel that playing the game gave you a sense of belonging to a community?’ ................................................................. 151
Table 8.17. Responses to the question ‘Do you feel that playing the game had an effect on your sense of loneliness?’ .............................................................................................. 151
Table 8.18. Responses to the question ‘Do you feel that playing the game had an effect on your sense of satisfaction with life?’ ................................................................. 152
Table 8.19. Compliance for participants who provided qualitative responses. .......... 159

List of figures

Figure 3.1. Pentagon copying in the Mini Mental State Examination: the original figure (Ridha and Rossor, 2005). .......................................................... 33
Figure 3.2. Pentagon copying in the Mini Mental State Examination: acceptable copies (Ridha and Rossor, 2005). .......................................................... 33
Figure 3.3. Pentagon copying in the Mini Mental State Examination: unacceptable copies (Ridha and Rossor, 2005). ......................................................... 33
Figure 3.4. Clock drawing examples for each of the severity scores from 0 to 5 (Shulman, 2000). ................................................................................... 35
Figure 3.5. Stroop word test (Chudler, 2015). ......................................................... 36
Figure 3.6. Stroop colour-word test: the task is to name the colour of each word, not to read the word (Chudler, 2015). ......................................................... 36
Figure 3.7. Trail-making tests A and B (Winner, 2009). ............................................. 38
Figure 3.8. Trail making test implemented in PEBL (Mueller, 2012). ....................... 38
Figure 3.9. An example of a letter cancellation test sheet (Canadian Partnership for Stroke Recovery, Letter cancellation test sheet). ........................................... 39
Figure 4.1. US Video game consumption vs violent crime, 1996-2004 (The Economist, 2005). ........................................................................................................ 63
Figure 5.1. Target Fun video game (Atari Inc., 1977) ................................................. 68
Figure 6.1. Administration web page ...................................................................... 79
Figure 6.2. The game client displayed a calendar to the participant, showing when to play the game and when to take the integrated tests. ......................... 79
Figure 6.3. Initial visual style ................................................................................... 82
Figure 6.4. Second visual style ................................................................................ 82
Figure 6.5. Final game design .................................................................................. 83
Figure 6.6. Avatar customisation screen, showing two different characters. .......... 85
Figure 6.7. Soft keyboard ....................................................................................... 85
Figure 6.8. Reaction time test (a) showing instructions, (b) period before stimulus, (c) showing stimulus, (d) after response from participant ....................... 87
Figure 6.9. Stroop colour test. The ‘green’ button is the correct choice in this case. .... 88
Figure 6.10. Stroop word test. The ‘red’ button is the correct choice in this case. ....... 88
Figure 6.11. Stroop colour-word test. The ‘blue’ button is the correct choice in this case. 88
Figure 6.12. Sequential number trail-making test (test ‘A’). ..................................... 89
Figure 6.13. Alternating number-letter trail-making test (test ‘B’). ......................... 89
Figure 6.14. Letter and symbol cancellation tests .................................................... 90
Figure 6.15. Client results screen .......................................................................... 91
Figure 6.16. Stroop test results for participant 13 ...................................................... 101
Figure 6.17. Stroop test results for participant 15 ...................................................... 101
Figure 6.18. Stroop test results for participant 22 ...................................................... 101
Figure 7.1. Stroop test results for participant 13 ...................................................... 103
Figure 7.2. Stroop test results for participant 15 ...................................................... 103
Figure 7.3. Stroop test results for participant 22 ...................................................... 103
Figure 7.4. Trail-making results for participant 13 .................................................... 103
Figure 7.5. Trail-making results for participant 15 .................................................... 103
Figure 7.6. Trail-making results for participant 22 .................................................... 103
Figure 8.1. ‘What’s been happening’ page. .................................................. 115
Figure 8.2. Player Jason’s home page. ...................................................... 115
Figure 8.3. Another player’s home page (Ella), as viewed by player Jason. 116
Figure 8.4. Guestbook page for player (Ella), as seen by player Jason. 116
Figure 8.5. ‘Messages sent’ page. .......................................................... 117
Figure 8.6. ‘Choose your trade’ screen. .................................................. 118
Figure 8.7. ‘Make a trade’ screen. .......................................................... 118
Figure 8.8. Example of ‘daily digest’ email ............................................. 120
Figure 8.9. Screenshot showing one example question from the De Jong Gierveld 6-point loneliness Scale implementation ................................................. 121
Figure 8.10. Satisfaction With Life Scale example question showing implementation with seven-item scale ......................................................... 122
Figure 8.11. Simpler Satisfaction With Life implementation with five-item scale. ........................................... 122
Figure 8.12. Total time spent logged in the system, for each participant. 129
Figure 8.13. Total number of login sessions, for all 16 participants, in each phase. 129
Figure 8.14. Total login duration, for all 16 participants, for each phase. 130
Figure 8.15. Session duration frequency. ................................................ 130
Figure 8.16. Loneliness scores for participant 13 (brain-injured). 133
Figure 8.17. Loneliness scores for participant 15 (brain-injured). 133
Figure 8.18. Loneliness scores for participant 23 (not brain-injured). 134
Figure 8.19. Loneliness scores for participant 27 (not brain-injured). 134
Figure 8.20. Loneliness scores for participant 36 (not brain-injured). 134
Figure 8.21. SWLS results for participant 13 (brain-injured). 136
Figure 8.22. SWLS results for participant 15 (brain-injured). 136
Figure 8.23. SWLS results for participant 23 (not brain-injured). 137
Figure 8.24. SWLS results for participant 27 (not brain-injured). 137
Figure 8.25. SWLS results for participant 36 (not brain-injured). 137
Figure 8.26. Stroop test results for participant 13 (brain-injured). 140
Figure 8.27. Stroop test results for participant 15 (brain-injured). 140
Figure 8.28. Stroop test results for participant 23 (not brain-injured). 141
Figure 8.29. Stroop test results for participant 27 (not brain-injured). 141
Figure 8.30. Stroop test results for participant 36 (not brain-injured). 141
Figure 8.31. Trail making results for participant 13 (brain-injured). 144
Figure 8.32. Trail making results for participant 15 (brain-injured). 144
Figure 8.33. Trail making test results for participant 23 (not brain-injured). 145
Figure 8.34. Trail making results for participant 27 (not brain-injured). 145
Figure 8.35. Trail making results for participant 36 (not brain-injured). 145
Figure 8.36. Reaction time for participant 13 (brain-injured). 147
Figure 8.37. Reaction time for participant 15 (brain-injured). 147
Figure 8.38. Reaction time for participant 23 (not brain-injured). 148
Figure 8.39. Reaction time for participant 27 (not brain-injured). 148
Figure 8.40. Reaction time for participant 36 (not brain-injured). 148
Figure 8.41. Correlation between initial loneliness score and study compliance. (Higher loneliness score means ‘more lonely’.)  r = -0.27 ............................................. 155
Declaration

Whilst registered as a candidate for the above degree, I have not been registered for any other research award. The results and conclusions embodied in this thesis are the work of the named candidate and have not been submitted for any other academic award.

Word count: 45 670
### Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>2D</td>
<td>Two-dimensional</td>
</tr>
<tr>
<td>3D</td>
<td>Three-dimensional</td>
</tr>
<tr>
<td>AAC</td>
<td>Augmented and alternative communication</td>
</tr>
<tr>
<td>ABI</td>
<td>Acquired brain injury</td>
</tr>
<tr>
<td>AT</td>
<td>Assistive technology</td>
</tr>
<tr>
<td>BCI</td>
<td>Brain computer interface</td>
</tr>
<tr>
<td>HCI</td>
<td>Human-computer interaction</td>
</tr>
<tr>
<td>LMIC</td>
<td>Low- and middle-income countries</td>
</tr>
<tr>
<td>MMO</td>
<td>Massively multi-player online game</td>
</tr>
<tr>
<td>MMORPG</td>
<td>Massively multi-player online role-playing game</td>
</tr>
<tr>
<td>MMSE</td>
<td>Mini Mental State Examination</td>
</tr>
<tr>
<td>PC</td>
<td>Personal computer</td>
</tr>
<tr>
<td>PEBL</td>
<td>Psychology Experiment Building Language</td>
</tr>
<tr>
<td>PSE</td>
<td>Photosensitive epilepsy</td>
</tr>
<tr>
<td>QOL</td>
<td>Quality of life</td>
</tr>
<tr>
<td>RCT</td>
<td>Randomised controlled trial</td>
</tr>
<tr>
<td>SSRD</td>
<td>Single-subject research design</td>
</tr>
<tr>
<td>SWLS</td>
<td>Satisfaction With Life Scale</td>
</tr>
<tr>
<td>TBI</td>
<td>Traumatic brain injury</td>
</tr>
<tr>
<td>TIA</td>
<td>Transient ischaemic attack</td>
</tr>
<tr>
<td>VE</td>
<td>Virtual environment</td>
</tr>
<tr>
<td>VR</td>
<td>Virtual reality</td>
</tr>
<tr>
<td>WHO</td>
<td>World Health Organization</td>
</tr>
<tr>
<td>XP</td>
<td>‘Extreme Programming’</td>
</tr>
</tbody>
</table>
Abstract

Every year, millions of people suffer a stroke, and millions more sustain a traumatic brain injury. This thesis proposes that playing multi-player online video games could provide a form of therapy for brain-injured people. The original contribution to knowledge comprises a conceptual framework for this claim, the design of a game to be used as a research instrument, and the findings of pilot studies, conducted with brain-injured participants.

An initial literature review led to the formation of the core proposal. In order to conduct experimental research with brain-injured participants, ethical approval was sought and obtained from the Faculty of Technology Ethics Committee at the University of Portsmouth.

The first phase of the study concentrated on the iterative development of a prototype online multi-player game, which encouraged cooperative, altruistic interaction. This research instrument included integrated cognitive tests.

The second phase of the research was to conduct pilot studies with brain-injured participants. The aim of these studies was to refine the experimental method and the software design, and to gather results to determine whether a larger research project would be warranted.

The first experiment was conducted over four weeks in 2013. Results from the cognitive tests did not show any improvement due to playing the video game, but methodological issues were discovered, and were used to refine the experimental protocol and software.

The second experiment was conducted over eight weeks in 2014. New tests integrated with the game software measured loneliness and satisfaction with life. These results showed promise for online multi-player games to have the potential to provide emotional and cognitive therapeutic benefit. It is argued that further research in this area is warranted, and recommendations are provided for such work.
Dissemination

Publications


Presentations


10 September 2013: Participant at the BCS HCI 2013 Doctoral Consortium, Brunel University, London.

26 June 2013: ‘Multi-player online video games for cognitive rehabilitation’, presentation, University of Portsmouth Faculty of Technology Research Day.

22 May 2013: ‘Multi-player online video games for cognitive rehabilitation’, presentation, University of Portsmouth Centre for Healthcare Modelling and Informatics Healthy Computing Day.


26 September 2012: ‘Multi-player online video games for cognitive rehabilitation’, research seminar presentation, University of Portsmouth School of Computing.


5 November 2010 ‘Brain-controlled video games for cognitive rehabilitation’. Poster presented at Health Development Forum Showcase, University of Portsmouth.

Conferences

31 May 2013: Poster presentation, UK Stroke Assembly 2013, Nottingham.

Acknowledgements

Many people have helped me over the last six years. Firstly, thanks and love to my wonderful family for their love, support and understanding.

To my supervisory team, Professor Jim Briggs, Dr Louise Turner, and Dr Alice Good, thank you. The late Dr Paul Gnanayutham inspired me to start this research. After Paul fell seriously ill, Louise took over as unofficial first supervisor, giving an enormous amount of support. Paul’s health grew worse, and Jim took over as first supervisor, with Alice becoming my third supervisor.

Thanks to all the study and play test participants, especially those from Different Strokes and Headway.

This is dedicated to those who survive brain injury, their families and friends, and all those who help them.

In memory of two friends, Paul and Mala.
1 Introduction

This research aims to help people who have an Acquired Brain Injury (ABI), by investigating the potential of multi-player online video games as a form of therapy. An ABI is defined as a sudden onset, non-degenerative injury to the brain occurring since birth (Beecham, Perkins, Snell, and Knapp, 2009), and is one of the most common causes of disability and death in adults (Feigin, Barker-Collo, Krishnamurthi, Theadom, and Starkey, 2010).

There are two main causes of ABI: stroke and traumatic injury (Feigin et al., 2010). A stroke is caused by the interruption to the normal flow of blood to the brain: the World Health Organization defines stroke as

‘Rapidly developing signs of focal (or global) disturbance of cerebral function, lasting longer than 24 hours (unless interrupted by death) with no apparent non-vascular cause’ (Aho, Harmsen, Hatano, Marquardsen, Smirnov, and Strasser, 1980).

A traumatic brain injury (TBI) is an ABI caused by trauma such as a blow to the head, an impact with a blunt object, or penetration by a sharp object: Menon et al. (2010) define a TBI as ‘an alteration in brain function, or other evidence of brain pathology, caused by an external force’. Road accidents, falls, sports injuries, and assaults are the most common causes of TBI (Hyder, Wunderlich, Puvanachandra, Gururaj, and Kobusingye, 2007).

The global incidence of stroke is nearly 17 million per year (Mozzafarian et al., 2015). In the UK, stroke is the biggest single cause of major disability (Mackay and Mensah, 2004, p. 50). The worldwide incidence of traumatic brain injury (TBI) is an estimated 10 million per year (Hyder et al. 2007). Combining these figures - and conservatively at that - over 25 million people are affected by brain injury annually – more than the population of Australia (Central Intelligence Agency, 2015).

These incidence figures are not the same everywhere in the world. The burden falls most heavily on low- and middle-income countries (LMIC) (Mozaffarian et al., 2015; Ma, Chan, and Carruthers, 2014). There is an increasing incidence of stroke in LMIC (Mozaffarian et al., 2015), and TBI, as more road users mean more road accidents. This means that rehabilitation services are less likely to be accessible to those ABI survivors in the very regions with the highest incidence (Hyder et al., 2007).

What happens to a person who survives a brain injury? Disabilities due to TBI include physical impairments, cognitive, communicative and swallowing disabilities. Motor speech and language disorders can hamper interaction with family and friends, leading to isolation
(Murdoch and Theodoros, 2001). There is a high frequency of depression and anxiety after an ABI (Mateer, 2005).

Cognitive rehabilitation aims to improve a person's function in areas relevant to their everyday lives (Cicerone et al., 2000). It is the process of regaining lost skills, or developing coping mechanisms to replace them. The goal is for a cognitively impaired person to function safely, productively and independently (Mateer, 2005). Mateer further notes that most successful interventions involve multiple, eclectic approaches. Recovery is gradual, and some patients continue to improve even years after their brain injury. The mechanism for this is neuroplasticity, the ability of neurons in the brain to make new connections and to reorganise existing ones (Dubin, 2002).

This thesis proposes online multi-player video games as a potential form of cognitive rehabilitation for brain-injured people. Video games have been used therapeutically for many years, in a wide variety of medical contexts (Griffiths, Kuss, and Ortiz de Gortari, 2013). Action games in particular have been shown to improve the abilities of players in aspects of vision, cognitive function, decision-making, and attention, and even to enhance the ability to learn, through the mechanism of enhancing neuroplasticity (Bavelier, Green, Pouget, and Schrater, 2012).

Massively multi-player online video games (MMOs) allow many players to interact with each other, and feature advanced, detailed multi-player worlds (Griffiths, Davies, and Chappell, 2004). In these games, players enjoy the social aspects and being able to help each other (Wang and Wang, 2008). In this thesis, I aim to show the potential of these games as a form of therapy, which could help alleviate feelings of isolation, and propose that this may lead to measurable cognitive improvements.

1.1 Motivations

My personal motivation is due to the work of Dr Paul Gnanayutham, who I first met in 2008 when I attended a guest lecture he gave. There I learned for the first time of locked-in syndrome, and how a brain-computer interface (BCI) could be used to enable people to communicate when they had been unable to do so. Paul’s work (Gnanayutham, Bloor, and Cockton, 2005) gave a voice to people who were trapped, and was a life-changing inspiration to me. I wanted to help Paul in his work, and to follow in his footsteps, if possible. I worked on maintaining Paul’s BCI software, and in 2010 began this research. At that time I knew I wanted to help brain-injured people, but did not know what form this would take.
Over the course of this research I have spoken with many TBI and stroke survivors. They have helped shape this research, and later, the study participants worked with me to seek experimental evidence for the proposal. I have learned that brain injury can strike any of us, at any time in our lives. One of the study participants, Mala, died during the course of the study. Paul suffered a series of strokes in 2012 and died on 10 October 2013. Brain injury is cruel and devastating.

1.2 Research questions and hypotheses

Following a review of the literature on brain injury, BCIs, loneliness, video game therapy, and massively multi-player online video games, the direction of my research took shape. As will be described in chapter 2, video games have a long history of being used therapeutically (Griffiths, Kuss, and Ortiz de Gortari, 2013). Multi-player online video games (MMOs) may have additional potential, as players report that the social aspects of these games, and the opportunity to help other players, are their favourite features (Griffiths, Davies, and Chappell, 2004). MMOs can provide other benefits: it may be that they develop social and emotional skills, foster collaboration and cooperation, and promote critical thinking, creativity, and problem-solving skills (Anderson, 2010). It is proposed in this thesis that there is a link from the social, emotional therapeutic benefits of MMOs, to counteracting the negative effects of isolation, to cognitive benefits that an improvement in quality of life may bring. Neuroplasticity, the mechanism for learning, growth and development (Pascual-Leone, Amedi, Fregni, and Merabet, 2005), is the process by which the proposed cognitive improvements would take place.

The research questions posed are these:

**Research Question 1:** Could playing a multi-player online video game provide a form of therapy for a brain-injured person? Would it improve cognitive function, emotional function, or both? Could any improvements be shown to be due to the social interactions fostered?

**Research Question 2:** If multi-player online video games can provide a cognitive therapeutic benefit, what are the essential features of these games that provide the benefit? Is it belonging to a community, playing cooperatively, behaving altruistically, etc.?

From research question 1, we can formulate hypotheses to be tested experimentally. Firstly we will test whether some cognitive skills can be positively affected by playing a multi-player online game, leading to an experimental design with the hypothesis:
**H1:** Multiplayer online video games contribute to an improvement in cognitive function.

Secondly, we will test the effect of the game on subjective emotional wellbeing. This holistic but rather vague term is divided into two specific measures: loneliness and satisfaction with life. These two qualities can be measured using existing instruments (reviewed in chapter 3), leading to an experimental design to test the following two hypotheses:

**H2:** Multiplayer online video games contribute to an improvement in subjective loneliness.

**H3:** Multiplayer online video games contribute to an improvement in subjective quality of life.

The second research question rests on the premise that multi-player online games do indeed provide the proposed benefits. This being the case, further experiments could determine the most beneficial attributes or nature of these games. In the first instance, the strategy is to use a simplified, bespoke game, which can be altered to test the effect of adding and removing features.

### 1.3 Research approach

The research approach was developed following a further literature survey. The overall approach is experimental, using a single-subject research design with a small group of brain-injured participants (n=3), and another, larger non-brain-injured group (n=13). The single-subject design is well suited to a study where the number of participants is small (Janosky, 2005).

A previous study (Malec, Jones, Rao, and Stubbs, 1984), which is more fully described in Chapter 5, was used as a starting point for the experimental design, and this has allowed some comparisons between the results of the current study and that previous one.

Participants were recruited from two UK brain injury support organisations, Headway (headway.org.uk), and Different Strokes (differentstrokes.co.uk). Headway’s stated mission is to promote understanding of brain injury, and to provide information, support and services to survivors, their families and carers. Different Strokes aims to support younger stroke survivors and their families. As part of my recruitment strategy, I visited, exhibited, and spoke at conferences, group meetings, and rehabilitation sessions. These conversations helped to guide the research and refine the software that has been developed.

In the experiments, the participants played an online multi-player video game, which contained integrated tests. In the first experiment, these measures were broad-based
cognitive tests. In the second experiment, these were expanded to include short questionnaires for loneliness (De Jong Gierveld and Van Tilburg, 2006) and satisfaction with life (Diener, Emmons, Larsen, and Griffin, 1985). The tests were taken every week, while the time spent in the game - the independent variable - was changed. This results in a time-series of data for each participant, for each different test.

1.4 Aims and objectives
This thesis proposes that playing online multi-player video games could provide a form of therapy for a brain-injured person. The aims of this research are to find out if this is the case, and to discover the nature of any such improvements. Should such improvements be found, a further aim would then be to find ways of increasing the effectiveness of this therapy.

In order to achieve these aims, the research project has the following objectives.

- From the literature, show a legitimate basis for the proposed therapeutic use of multi-player online video games;
- Establish a research approach for finding evidence for the proposed benefits;
- Develop game and research software as required;
- Conduct experiments and collect data; and
- Disseminate the results in order to stimulate further research in this area.

1.5 Original contribution to knowledge
In this research project, experiments have been conducted with brain-injured and non-brain-injured participants, with the aim of finding measurable improvements in emotional and cognitive wellbeing that can be ascribed to playing an online multi-player game. The experimental design and the findings are an original contribution to knowledge. It is recognised that the number of participants in these experiments is low, and so the external validity of these results should not be exaggerated. Nevertheless, in terms of impact, the context is of increasing prevalence of brain injury worldwide. Many people could benefit from a new form of therapy.

1.6 Structure of the remainder of this dissertation
Chapter 2 reviews the literature on ABI, rehabilitation, and quality of life for people with an ABI. The therapeutic uses to which video games have been applied are also reviewed. These two topics are brought together to form the central thesis, that a multi-player online video game could provide a measurable therapeutic benefit to people with brain injury.
Chapter 3 provides a review of subject areas that are relevant to developing an approach to test the hypothesis. Usability, accessibility and designing for disabled people are reviewed. Research methodologies are surveyed, to find a suitable experimental method. The single-subject (also called single-case, or ‘n=1’) design is examined. Software engineering methodologies, human-computer interaction, and video game design principles are also reviewed.

Chapter 4 surveys the ethical and medical considerations that apply when working with disabled participants. The potentially negative effects of playing video games are described, and the issues of consent and privacy reviewed.

Following this survey, the research approach is presented in chapter 5, detailing the experimental design. The design is an extension of a prior study (Malec et al., 1984), which is described in some detail.

A major part of the preparation for the experiments was the development of a software system: a prototype multi-player online video game and associated cognitive tests. This encompassed client and server-side software, for the game and tests, data storage and reporting. Chapter 6 describes the design of this software system.

Chapter 7 describes the experiment conducted over four weeks during August-September 2013. This experiment focused on three brain-injured participants, and aimed to measure any improvements in cognitive test performance, which could be ascribed to playing the prototype multi-player online video game.

Chapter 8 describes the 8-week experiment of May-July 2014. For this experiment, the software was redeveloped to provide a social interactive experience, when the players may not be online at the same time. The cognitive tests were extended to include measures of loneliness and satisfaction with life.

Chapter 9 summarises the work and results, and evaluates the contribution to the video game health research community. We conclude with a discussion of future work that could be undertaken in the study area.
2 Background and research questions

This chapter reviews the literature on acquired brain injury (ABI), rehabilitation, and quality of life for people with an ABI. The therapeutic uses to which video games have been applied are also reviewed. These two topics are brought together to form the basis for the proposal that a multi-player online video game could provide a measurable therapeutic benefit to people with an ABI.

2.1 Acquired Brain Injury (ABI)

ABI is defined as a sudden onset, non-degenerative injury to the brain occurring since birth (Beecham, Perkins, Snell, and Knapp, 2009), and is one of the most common causes of disability and death in adults. The leading causes of acquired brain injury are stroke and traumatic injury (Feigin et al., 2010).

2.1.1 Stroke

A cerebrovascular accident, commonly called a stroke, is an ABI caused by the interruption to the normal flow of blood to the brain. The word ‘stroke’ was probably first used by physician William Cole, in 1689 – prior to this, ‘apoplexy’ was the term used to describe acute brain injury of a non-traumatic nature (Sacco et al., 2013). A 1980 World Health Organization report defined stroke as ‘rapidly developed clinical signs of focal (or global) disturbance of cerebral function, lasting more than 24 hours or leading to death, with no apparent cause other than of vascular origin’ (Aho, Harmsen, Hatano, Marquardsen, Smirnov, and Strasser, 1980). A more comprehensive and updated definition has been proposed by the American Heart Association and American Stroke Association (Sacco et al., 2013), shown in Table 2.1 (overleaf).

There are three main types of stroke (Centers for Disease Control and Prevention, 2013). The most common form is ischaemic, in which a narrowing or blockage of arteries restricts blood and oxygen flow to the brain. The blockages are often formed of blood clots caused by arterial plaque. The second, and less common form of stroke is haemorrhagic, where an artery leaks or bursts. The resulting haemorrhage damages brain cells by exerting pressure on them. The third, and least common form of stroke is a subarachnoid haemorrhage, bleeding on the surface of the brain. Additionally, a transient ischaemic attack (TIA) or ‘mini stroke’ is caused by a temporary disruption in the blood supply to part of the brain.
Table 2.1. Definition of stroke (Sacco et al., 2013).

<table>
<thead>
<tr>
<th>Definition of stroke</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Definition of CNS infarction</strong>:</td>
<td>CNS infarction is brain, spinal cord, or retinal cell death attributable to ischemia, based on</td>
</tr>
<tr>
<td>1. pathological, imaging, or other objective evidence of cerebral, spinal cord, or retinal focal ischemic injury in a defined vascular distribution; or</td>
<td></td>
</tr>
<tr>
<td>2. clinical evidence of cerebral, spinal cord, or retinal focal ischemic injury based on symptoms persisting ≥24 hours or until death, and other etiologies excluded.</td>
<td></td>
</tr>
<tr>
<td>(Note: CNS infarction includes hemorrhagic infarctions, types I and II; see “Hemorrhagic Infarction.”)</td>
<td></td>
</tr>
<tr>
<td><strong>Definition of ischemic stroke</strong>:</td>
<td>An episode of neurological dysfunction caused by focal cerebral, spinal, or retinal infarction.</td>
</tr>
<tr>
<td>(Note: Evidence of CNS infarction is defined above.)</td>
<td></td>
</tr>
<tr>
<td><strong>Definition of silent CNS infarction</strong>:</td>
<td>Imaging or neuropathological evidence of CNS infarction, without a history of acute neurological dysfunction attributable to the lesion.</td>
</tr>
<tr>
<td><strong>Definition of intracerebral hemorrhage</strong>:</td>
<td>A focal collection of blood within the brain parenchyma or ventricular system that is not caused by trauma.</td>
</tr>
<tr>
<td>(Note: Intracerebral hemorrhage includes parenchymal hemorrhages after CNS infarction, types I and II—see “Hemorrhagic Infarction.”)</td>
<td></td>
</tr>
<tr>
<td><strong>Definition of stroke caused by intracerebral hemorrhage</strong>:</td>
<td>Rapidly developing clinical signs of neurological dysfunction attributable to a focal collection of blood within the brain parenchyma or ventricular system that is not caused by trauma.</td>
</tr>
<tr>
<td><strong>Definition of silent cerebral hemorrhage</strong>:</td>
<td>A focal collection of chronic blood products within the brain parenchyma, subarachnoid space, or ventricular system on neuroimaging or neuropathological examination that is not caused by trauma and without a history of acute neurological dysfunction attributable to the lesion.</td>
</tr>
<tr>
<td><strong>Definition of subarachnoid hemorrhage</strong>:</td>
<td>Bleeding into the subarachnoid space (the space between the arachnoid membrane and the pia mater of the brain or spinal cord).</td>
</tr>
<tr>
<td><strong>Definition of stroke caused by subarachnoid hemorrhage</strong>:</td>
<td>Rapidly developing signs of neurological dysfunction and/or headache because of bleeding into the subarachnoid space (the space between the arachnoid membrane and the pia mater of the brain or spinal cord), which is not caused by trauma.</td>
</tr>
<tr>
<td><strong>Definition of stroke caused by cerebral venous thrombosis</strong>:</td>
<td>Infarction or hemorrhage in the brain, spinal cord, or retina because of thrombosis of a cerebral venous structure. Symptoms or signs caused by reversible edema without infarction or hemorrhage do not qualify as stroke.</td>
</tr>
<tr>
<td><strong>Definition of stroke, not otherwise specified</strong>:</td>
<td>An episode of acute neurological dysfunction presumed to be caused by ischemia or hemorrhage, persisting ≥24 hours or until death, but without sufficient evidence to be classified as one of the above.</td>
</tr>
</tbody>
</table>

CNS indicates central nervous system.

The most important risk factor for stroke is age (Stroke Association, 2016). Other risk factors for stroke include the following (Mozaffarian et al., 2015; Feigin et al., 2010):

- High blood pressure (hypertension),
- Existing conditions: diabetes mellitus, and disorders of heart rhythm,
- High blood cholesterol and other lipids, leading to atherosclerosis,
- Smoking,
- Physical inactivity,
• Nutrition, and
• Family history.

Alcohol consumption is related to stroke in that it is associated with hypertension and atrial fibrillation (slow, abnormal heart rhythm), but a more direct relationship between alcohol and stroke is unclear (Feigin et al., 2010).

Some of these risk factors are manageable by lifestyle changes or medication, and there has indeed been a reduction in stroke incidence in high-income countries between 1990 and 2010 (Mozaffarian et al., 2015). Hypertension can be managed with medication. Having conducted a meta-study encompassing 19 trials, including nearly 45,000 participants, Xie et al. (2016) recommend lowering the threshold at which blood pressure-lowering medication is prescribed, to reduce the incidence of strokes further.

The increasing incidence of strokes in low- and middle-income countries has offset improvements in high-income countries, and the absolute number of people who have strokes annually has increased over the period 1990-2010. Worldwide, there were an estimated 17 million events of stroke in 2010 (Mozaffarian et al., 2015). In the UK, there are over a million stroke survivors (Townsend et al., 2012, pp. 57-58): stroke is a leading cause of serious long-term disability (Mozaffarian et al., 2015).

2.1.2 Traumatic Brain Injury (TBI)

A traumatic brain injury (TBI) is defined as an alteration in brain function, or other evidence of brain pathology, caused by an external force (Menon, Schwab, Wright, and Maas, 2010). Common causes of TBI are motor vehicle accidents, bicycle accidents, assaults, falls, and sports injuries (Hyder, Wunderlich, Puvanachandra, Gururaj, and Kobusingye, 2007; Lindsay and Bone, 2004, p. 216; Ponsford, Sloan, and Snow, 2001). Motor vehicle accidents are the most common cause of TBI globally (Hyder et al., 2007), although for children this may be as a pedestrian or cyclist rather than passenger (Murdoch and Theodoros, 2001).

There are two main types of TBI: open injury, where the skull is penetrated, and closed injury. Closed head injuries are much more common for civilians, although penetrating injuries are more common in wartime (Murdoch and Theodoros, 2001). In a closed TBI, the brain is subjected to compression, acceleration, and rotational forces, causing brain tissue to be torn and sheared. The head may be crushed, but more commonly the victim suffers a brief impact, causing a sudden movement of the head. The brain may collide with the skull at the point of impact. Another injury opposite this point may then occur as the brain rebounds (Murdoch and Theodoros, 2001).
The primary mechanism in many cases of closed TBI is diffuse axonal injury (DAI). DAI is the term given to widespread damage caused by the shearing or rotational forces (Ponsford, Sloan, and Snow, 2001). At the microscopic level, the direction of the shear may be visible (Lindsay and Bone, 2004, p. 218). DAI may be widely distributed, and occur deep in the brain, in the white matter and brain stem (Murdoch and Theodoros, 2001). Even a minor head injury where the patient loses consciousness results in some damage to neurons. As the ability of these cells to regenerate is limited, repeated head injuries have a cumulative effect (Lindsay and Bone, 2004, p. 218).

TBI has two stages: primary and secondary. The primary stage is the initial trauma at the moment of impact, such as DAI, contusions, lacerations, haemorrhage, and lesions (Murdoch and Theodoros, 2001). Secondary brain injury describes complications that may follow, which are potentially treatable. The main cause of secondary brain damage is cerebral ischaemia. Other causes include haematoma, brain swelling, infection, raised intracranial pressure, respiratory failure, and hypotension (Ponsford, Sloan, and Snow, 2001). Murdoch and Theodoros (2001) also describe damage due to herniation, cerebral atrophy, and ventricular enlargement.

Hyder et al. (2007) estimate that 10 million people per year are affected by TBI globally. TBI is especially prevalent in low- and middle-income countries (LMIC), where the risk factors tend to be higher, and health systems are inadequately prepared. Road traffic injuries in Latin America and Sub-Saharan Africa are the main cause of a higher TBI incidence in these regions.

Improvements in road safety have reduced the number of people who suffer a brain injury. Cook and Sheikh (2000) report a 12% reduction in cyclist head injuries in England between 1991 and 1995, ascribed to the increased use of bicycle helmets over the period. Reductions in drink-driving and increased use of seat belts, crash helmets and airbags have reduced the incidence of head injury in many countries (Lindsay and Bone, 2004, p. 216).

2.1.3 Effects of brain injury

Different regions of the brain have different functions, and impairments suffered by an injured person will therefore depend on which areas of the brain were damaged. Despite the uniqueness of every brain injury, there are common symptoms. Powell (1994) lists the effects of brain injury most often noted by relatives of a brain-injured person. These effects include personality changes, slowness, poor memory, irritability, bad temper, tiredness, depression, rapid mood changes, tension and anxiety, and threats of violence. Murdoch and Theodoros (2001, p. ix) note that disabilities due to TBI typically include physical
impairments, cognitive, communicative and swallowing disabilities, and that motor speech and language disorders hamper interaction with family and friends, leading to isolation. There is a high frequency of depression and anxiety after an ABI (Mateer, 2005). Lerdal et al. (2009) state that fatigue is a common complaint following a stroke.

### 2.1.4 Assessment of brain injury

When a person suffers a moderate or severe brain injury, he or she will enter a comatose state (Powell, 1994). The duration of the coma is an indicator of the severity of the injury, and it is also possible to assess the severity of the injury during the comatose period, by gauging the responsiveness of the patient.

The Glasgow Coma Scale (Teasdale and Jennett, 1974) is used by clinical staff to measure the status of a head injury patient and provides a basic assessment of their chances of survival. The patient’s ability to open his or her eyes, verbalise, and move is rated, giving an overall number from 3 (most severe impairment) to 15 (least impairment) (Lindsay and Bone, 2004). A rating of 8 or less indicates that the patient is comatose (Ponsford, Sloan, and Snow, 2001).

The extent of a brain injury can be revealed using imaging technologies. Computerised Axial Tomography (CT or CAT) scanning is an X-ray based technique, used for detecting foreign bodies, tumours, etc. Because CT scans expose patients to radiation, a blood test to diagnose brain injury could be a preferable alternative in future. Concentrations in the blood of signature proteins correspond to brain injury severity, though at such low levels that they are currently difficult to measure (Orcutt, 2015).

Magnetic Resonance Imaging (MRI) uses magnetic fields and radio signals, and resolves soft tissues well. Functional MRI (fMRI) detects changes in localised activity, relying on changes in blood oxygen concentration in active areas of the brain (Dubin, 2002). Single Photon Emission Computed Tomography (SPECT) and Positron Emission Tomography (PET) also show brain function (Lindsay and Bone, 2004).

Upon regaining consciousness, the patient will experience a period of post-traumatic amnesia (PTA). The period of PTA is judged to have ended when the patient is able to form new memories (Ponsford, Sloan, and Snow, 2001). The periods of the coma and of the PTA give a reliable indication of the severity of the brain injury. A coma period of more than six hours, or PTA of more than 24 hours is classed as a severe injury; such cases account for 5% of all head injuries (Powell, 1994).

Some patients remain in the comatose state, or transition to a persistent vegetative state (PVS). PVS patients are unable to move or communicate, and are not aware. Some other
patients are cognitively intact and aware of their surroundings, but are unable to move or communicate, a condition known as locked-in syndrome. Cases have occurred of patients who were misdiagnosed as being in PVS, when they were in fact locked in. Monti et al. (2010) describe patients who were outwardly non-aware and non-communicative, but who could answer questions using MRI scanning. As patients diagnosed as PVS are more routinely scanned for cognitive activity, so the number of diagnosed locked-in patients may increase, and the number of PVS patients decrease correspondingly (Monti et al., 2010).

2.2 Rehabilitation and quality of life after brain injury

This section is concerned with rehabilitation after brain injury. Johnson and Rose (1996, p. 185) note that although every brain injury is unique, and the outcome uncertain, an attempt must be made to predict and maximise outcome after brain injury.

2.2.1 Cognitive rehabilitation

Cicerone et al. (2000) define cognitive rehabilitation as a ‘systematic, functionally oriented service of therapeutic activities that is based on assessment and understanding of the patient’s brain-behavioural deficits’. They stress that the aim of cognitive rehabilitation services is to improve a person’s function in areas relevant to their everyday lives. For people who have suffered a brain injury, rehabilitation is the process of regaining lost skills, or developing coping mechanisms to replace them. The goal is for a cognitively impaired person to function safely, productively and independently (Mateer, 2005).

Rehabilitation has two stages. The first is the acute stage, where medical professionals stabilise the patient. The second stage is where family and carers take over (Powell, 1994). The process of recovery for an individual is affected by multiple variables, only one of which is the severity of the brain injury (Sohlberg and Mateer, 2001). Every patient responds differently to treatment, and different skills may be regained at different times (Sohlberg and Mateer, 2001, p. 69). Some patients continue to improve even years after the brain injury. For example, dysarthria (characterised by slurred, indistinct speech) varies widely in severity and recovery, with some patients showing ‘significant and continuous’ improvement (Theodoros, Murdoch, and Goozée, 2001).

Mateer (2005) highlights some general principles for successful cognitive rehabilitation. These are the need to take an individualised approach, to involve clients and caregivers in all aspects, and clear and realistic goal setting. Mateer further notes that most successful interventions involve multiple, eclectic approaches. According to Sohlberg and Mateer (2001, p. 10), a basic assumption underlying cognitive rehabilitation is that cognition
cannot be treated in isolation: ‘Brain damage affects cognitive, social, behavioural and emotional functioning’.

Ben-Yishay and Daniels-Zide (2000) discuss evidence that optimal outcomes after rehabilitation require that the individual achieves an ‘examined self’. These authors suggest that acceptance of the disability – as a positive, active process, rather than resignation – 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 and Mateer (2001) to mean 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, although this would not be the sole focus.

### 2.2.2 Cognitive remapping (neuroplasticity)

At the anatomical level, cognitive remapping or neuroplasticity is the ability of neurons in the brain to make new connections and to reorganise existing ones (Dubin, 2002). More plasticity is available at birth than later in life, but some plasticity within the adult cortex does occur, based on the sprouting of new axonal branches. For example, musicians develop a larger amount of sensory cortex devoted to their fingers, which require greater speed and precision. Dubin (2002) concludes that plastic changes in the cortex are ongoing throughout adulthood, suggesting that carefully designed rehabilitation exercises might reduce the effects of ABI. Pascual-Leone, Amedi, Fregni, and Merabet (2005) state that plasticity is an intrinsic property of the nervous system, retained throughout the lifespan. Similarly, studies by Hummel et al. (2005) have produced evidence using functional imaging that neuronal tissues undergo plastic changes, in healthy and brain-injured people. More recently it is has become evident that the extent of neuroplastic change can be augmented (Vines, Nair, and Schlaug, 2008).

Gage (2004) reports that a stroke triggers a notable increase in the production of new cells, a process called neurogenesis. In a large stroke such ‘micro-repair’ is inadequate, but may protect and repair the brain after small, unrecognised strokes. Because the brain of a young person has greater plasticity, age is one of the main considerations in predicting outcome after brain injury (Johnson and Rose, 1996, p. 186).

### 2.2.3 Quality of life

The World Health Organization defines quality of life (QOL) as

‘Individuals’ perception of their position in life in the context of the culture and value systems in which they live and in relation to their goals, expectations,
standards and concerns. It is a broad-ranging concept affected in a complex way by the person’s physical health, psychological state, level of independence, social relationships, personal beliefs and their relationship to salient features of their environment’ (World Health Organization, 1997).

According to Diener and Suh (1997) there are three major philosophical approaches to determining quality of life. These are:

- Normative ideals based on religious, philosophical, or other systems;
- The satisfaction of material preferences; and
- The subjective experience of the individual.

The first two approaches are ostensibly objective; the third is based on reported, subjective feelings of joy, pleasure, contentment, and life satisfaction. Subjective quality of life does not depend on income, age, or sex (Myers and Diener, 1995).

Rapley (2003) notes that the term ‘quality of life’ is used inconsistently – even with ‘abandon’. QOL is defined in a wide variety of ways, but studies routinely lack a formal definition of QOL. Additionally, widely used measures of QOL fail to relate to an explicit theory of QOL (Rapley, 2003, p. 29). The difficulty with measuring QOL is that it is unique to each individual. Many standardised measures, which are widely used, can only capture the QOL of individuals in a limited way (Carr and Higginson, 2003).

Bowling (2003) states that QOL consists of multiple subjective and objective dimensions, and is dynamic, changing over time. Carr and Higginson (2003) argue that QOL measures should be patient-centred rather than standardised, but analysis, interpretation and comparison of individualised measures is of course more difficult. Seale and Turner-Smith (2003) report measures that have been used by rehabilitation professionals to measure the effect of assistive technology (AT) on QOL. These include health-related QOL instruments, participation-oriented instruments, and AT-specific measures.

Jones, Haslam, Jetten, Williams, Morris, and Saroyan (2010) developed the Trauma and Recovery Experiences Assessment Tool (‘TREAT’) questionnaire to measure quality of life for people with ABI. This questionnaire asks respondents about their own reaction to the injury, social reactions, social support, and the impact of the injury on personal relationships. Jones and team assessed 630 individuals with an ABI, and found a surprising positive relationship between injury severity and life satisfaction. These 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’.

2.2.4 The disability paradox

The so-called ‘disability paradox’ is that individuals who experience severe difficulty performing everyday tasks may still report good or excellent QOL (Carr and Higginson, 2003). Albrecht and Devlieger (1999) conducted semi-structured interviews with 153 disabled people. Over half (54.3%) reported excellent or good quality of life. This is at odds with the widespread perception that disabled people do not have as high a quality of life as non-disabled people (Albrecht and Devlieger, 1999). That some people with disabilities have a deep sense of wellbeing leads these authors to conclude that disability calls accepted values and notions of wellbeing into question.

Myers and Diener (1995) offer a possible explanation for the disability paradox. Over time, the immediate response to significant life events fades. Even the trauma of a paralysing accident typically gives way to the return of normal levels of happiness. The more recent an event, the greater its effect; but only events which happened within the last three months influence subjective wellbeing. Even so, a person who is in a nursing home may have a feeling that they have no personal control, and so this may affect their subjective happiness.

2.2.5 Loneliness

This section explores the effects of loneliness, and how it relates to ABI. Isolation is a common consequence of ABI (Murdoch and Theodoros, 2001), and the effects of loneliness can contribute to cognitive and emotional decline, among a host of other negative symptoms. Hawkley and Cacioppo (2010) define loneliness as ‘a distressing feeling that accompanies the perception that one’s social needs are not being met by the quantity or especially the quality of one’s social relationships’. Other definitions also rest on the perception of the sufferer (Pinquart and Sörensen, 2001): loneliness is subjective, the measure of a person’s perceived isolation.

In his 1973 work, Weiss found that loneliness is distinct from ‘aloneness’ or physical isolation; and is neither a simple desire for company, nor grief, nor depression, although depression and loneliness can go hand in hand (Hawkley and Cacioppo, 2010). Weiss identified two different types of loneliness: emotional and social. Emotional loneliness forms in the absence of a close emotional attachment, and ‘can only be remedied by the integration of another emotional attachment or the reintegration of the one that had been lost’. Social loneliness is associated with the absence of an engaging social network. These
distinct types of loneliness have different symptoms: emotional isolation has symptoms of anxiety and emptiness; for social isolation, the symptoms are feelings of boredom, aimlessness, and marginality.

Loneliness can have dramatic physical effects on the sufferer. Hawkley and Cacioppo (2010) report negative effects of loneliness including increased risk for cardiovascular mortality, and an increased rate of physiological ageing. Emotionally and cognitively, the effects are no less severe, with reported impairments in cognitive ability and executive function, and association with psychosis, suicide, depression, and increased risk of Alzheimer’s disease.

Loneliness may affect ABI survivors in particular, as physical, cognitive, and communicative impairments hinder interaction with others. This isolation is not Weiss’ ‘aloneness’, although restrictions to mobility could indeed cause physical isolation; but rather, would perhaps be best described as social loneliness, where the sufferer is marginalised due to his or her disabilities. ABI has been described as a ‘silent epidemic’ (Hyder et al., 2007), as many survivors are not obviously disabled by their injury: a physically able ABI survivor could still find himself or herself socially excluded.

2.2.6 Summary

To conclude this section on rehabilitation and quality of life after brain injury, the points most relevant to the formation of the hypothesis are summarised.

Every year, over 15 million people suffer a stroke, and this number is rising. Stroke is the biggest single cause of major disability in the UK, and a major cause of disability worldwide. Additionally, TBI affects some ten million people. TBI is most prevalent in low- to middle-income countries, where access to rehabilitation services can be limited. Many people with ABI are isolated, and the subjective perception of isolation – i.e. loneliness – can have a dramatic effect on the sufferer, physically, emotionally, and cognitively. Rehabilitation seeks to provide greater independence, and to improve the cognitive and emotional wellbeing of the patient. Neuroplasticity is the mechanism by which the brain can repair, or adapt, over time. Long term, the quality of life for a person with ABI may be good or excellent, but this may depend on the richness of his or her social networks and relationships.

The next section surveys the literature on the many ways that video games have been used therapeutically. One particular kind of video game, massively multi-player online games, may have the potential to alleviate the isolation of ABI sufferers that has been described.
2.3 Therapeutic uses of video games

Video games are not just entertaining; they have been used for medical rehabilitation. Griffiths, Kuss, and Ortiz de Gortari (2013) summarise many ways in which video games have been used for therapeutic treatment, stating

‘Research dating right back to the early 1980s has consistently shown that playing computer games (irrespective of genre) produces increases in reaction times, improved hand-eye coordination, and raises players’ self-esteem. What’s more, curiosity, fun, and the nature of the challenge also appear to add to a game’s therapeutic potential’ (Griffiths, Kuss, and Ortiz de Gortari, 2013).

2.3.1 Flow, play and games

Video games are new, but games are not – board games have been unearthed dating from the Neolithic period (Flanagan, 2009). Landmark analyses of play and games by Huizinga (1949) and Caillios (1961) have noted the importance of play to people’s lives. Huizinga defined play as follows:

‘Play is a free activity standing quite consciously outside ‘ordinary’ life as being ‘not serious,’ but at the same time absorbing the player intensely and utterly. It is an activity connected with no material interest, and no profit can be gained by it. It proceeds within its own proper boundaries of time and space according to fixed rules and in an orderly manner.’

Caillios defined play as an activity having these characteristics:

- Free, as in not obligatory: once made compulsory, an activity ceases to be play;
- Separate from reality, existing within what Huizinga termed the ‘magic circle’;
- Having a result which is uncertain;
- Unproductive, in that no goods or wealth are created;
- Governed by rules; and
- Make-believe: the players are aware that this is not real life.

Both Caillios and Huizinga stressed that a game must be played voluntarily. Professional sportsmen and women are not playing, but working; for a professional actor, a performance is not the game of mimicry, but working to provide a simulation. To this we could add that a person who is working at testing a video game is not at play; and neither would participants in a video game study, if they did not feel that they were playing entirely voluntarily. Caillios notes that even in games that are entirely personal, competitiveness with others can easily materialise, and that games generally presuppose company, not solitude.
The term *autotelic* describes an activity that is performed not with the expectation of some future benefit, but for the reward of simply doing it. In seeking to optimise human happiness, Csikszentmihályi (2002) defined ‘flow’ as an autotelic experience: the way people describe their state of mind when consciousness is ‘harmoniously ordered’. For Csikszentmihályi, ‘Games are an obvious source of flow, and play is the flow experience par excellence’.

### 2.3.2 Video games

Video games are a new medium and – surely – art form, born in the second half of the twentieth century. The term ‘arcade game’ refers to stand-alone devices, played in public places such as amusement arcades (Raessens and Goldstein, 2005). The term ‘console game’ refers to games that are played on a dedicated hardware device, called a console, connected to a television. ‘Handheld games’ are portable game devices, the current generation of which primarily consists of the Sony Playstation Vita® and Nintendo DS® range. In fact this generation of dedicated handheld gaming devices may well be the last, as they are overtaken by ever more powerful and numerous smartphones (Brightman, 2015). Finally, ‘computer games’ are entertainment software applications played on a personal computer (Raessens and Goldstein, 2005). The term ‘video game’ is used here to mean a computer game, handheld game or console game – video game arcades are a rarity now, sadly.

The term ‘serious game’ was originally defined to mean a game designed for an educational purpose, rather than primarily for amusement (Abt, 1970, p. 9). In the context of video games, a serious game is one whose primary purpose is not entertainment, such as an educational or training game, or a game designed to promote rehabilitation. Other serious games have been designed to effect social change, addressing issues of poverty, racism and other forms of discrimination, war, and human rights (Flanagan, 2009, p. 243). Viewed as software, video games have some unique characteristics, identified by Neal (1990). Games are used for their own sake, whereas other types of software are invariably tools used as a means to an end. Games typically have clear goals and objective measures of success, such as a score; but game players tend to also create their own subjective measures of success. Playing a video game requires some degree of motor control, which may be demanding for real-time games where speed is a factor. Additionally, to succeed at a game the player needs to understand its rules and develop strategies. This may be required to determine the optimal next move, or placement of a game entity; or a search strategy may need to be employed. Success at a game may require curiosity on the part of
the player, should the game world need to be explored to find new locations or other game entities.

2.3.3 Video games for physical and behavioural therapy

This section reviews examples of contexts in which video games have been used for therapy.

Games have been successfully used in situations where repetitive motion is required of a patient, as physiotherapy. This success may be due to the motivating nature of games, and their role in distracting attention from discomfort during physiotherapy (Griffiths, Kuss, and Ortiz de Gortari, 2013). O’Connor and team (2000) developed a wheelchair interface to computer games called Game Wheels. This interface motivated spinal cord injury patients to exercise more regularly, by controlling games by driving their wheelchair. Sietsema, Nelson, Mulder, Mervau-Scheidel, and White (1993) report the use of an electronic table-top game to promote arm reach for people who have suffered a TBI. Betker, Desai, Nett, Kapadia, and Szturm (2007) used bespoke serious games to promote balance for people with spinal cord and traumatic brain injuries. Burke and colleagues (Burke, McNeill, Charles, Morrow, Crosbie, and McDonough, 2009) developed a suite of serious games designed to aid recuperation after a stroke. This was accomplished by requiring the player to make repetitive arm movements that aid upper limb recovery.

In a 2006 survey of game players, commissioned by game development company PopCap, it emerged that twenty-seven percent of the game players claimed that distraction from pain was a benefit (DeMaria, 2007, p. 34). In fact video games have been used to distract sufferers from pain for many years. Video games have been found to be an effective way to increase sitting tolerance for lower back pain sufferers (Butler, 1985), and Adriaenssens, Eggermont, Pyck, Boeckx, and Gilles (1988) report their therapeutic use of video games for burned children. Rehabilitation programs for burn patients focus on maintaining mobility. The exercises may be unpleasant, and young children may not understand the need for them. Some exercises such as clenching the fist and moving the wrist are particularly difficult to perform. Video games can be used to encourage such hand movements, using specially designed controllers. As the motions control a game, the child receives immediate reward for the effort – even children as young as two years old responding. The authors stress that the use of video games does not replace the physiotherapist, but is a complementary therapy.

Video games have been used as therapy for children for many years, in medical contexts including children with learning disabilities (Loftus and Loftus, 1983, p. 148), autism
(Demarest, 2000), severe psychiatric disorders (Favelle, 1994), and emotional and behavioural problems (Spence, 1988).

2.3.4 Video games for cognitive therapy

This section looks specifically at examples where video games have been used in a cognitive therapeutic context, rather than for physical or behavioural benefit.

Malec, Jones, Rao, and Stubbs (1984) performed early research into the use of video games as cognitive therapy, with a small group of ten ABI participants. The results showed no improvement in cognitive abilities, above those that may occur spontaneously, except for an improvement in reaction time. This study is described in more detail in chapter 5.

Fisher (1986) proposed that computers could be useful tools in cognitive rehabilitation, as a way of measuring cognitive skill and improvement, as a means of providing therapy, and a way of providing motivation. Noting that game software should be matched to a patient’s interests as well as disabilities, Fisher describes some serious games that have been developed as therapy for perceptual disorders, conceptual thinking, attention, concentration and memory, and language difficulties. She concludes that computer games may make the struggle of rehabilitation easier, by providing some crucial motivation.

Playing action video games may provide a general improvement in the ability to learn. Bavelier, Green, Pouget, and Schrater (2012) review many cases where players of action video games have improved abilities in some aspects of vision, cognitive function, decision-making and attention. Playing such games may enhance the ability to learn, through the mechanism of enhancing neuroplasticity. This is in contrast to other training scenarios where, rather than a general improvement, only an improvement in the specific training task tends to be found.

2.3.5 Multi-player and massively multi-player online games

Video games, from the very earliest examples, have been played by more than one player (Winter, 2013; Malliet and de Meyer, 2005). Massively multi-player online games (MMOs) allow many players to interact with each other, and feature advanced, detailed multi-player worlds (Griffiths, Davies, and Chappell, 2004). Popular massively multi-player online role-playing games (MMORPGs) have millions of players worldwide – e.g. World of Warcraft (Blizzard Entertainment, 2004-2015) had 12 million players at its peak in 2010 (Purchese, 2015). The actual number of players in any part of the game world at any one time is far fewer, due to limitations of bandwidth and computation speed of the server on which the game world is running.
Griffiths, Davies and Chappell (2004) report a study of 540 Everquest (Daybreak Game Company, 1999-2015) players. The favourite features of this online game reported by the players were the social aspects of playing online, such as social contact, and helping others. In this study, 81% of players were male, with a mean age of 28 years. Wang and Wang (2008) studied pro-social behaviour and gender in a Taiwanese online game. They found that the participants help each other, engaging in altruism and reciprocity. Male players were found to prefer helping females rather than other males, leading Wang and Wang to conclude, fairly unsurprisingly, that male gamers desire friendship with the opposite sex. Indeed, Yee (2003) found that approximately 10% of male players, and 33% of female players, have virtually married within an MMORPG.

Anderson (2010) presents advantages and limitations of MMORPGs in a learning context, noting that they are potentially psychologically therapeutic. Other advantages are that motivation is inspired, social and emotional skills are developed, collaboration and cooperation is fostered, and critical thinking, creativity, and problem-solving skills are promoted. Anderson also found that MMORPG players view joining a guild (an in-game player organisation) as preferable or necessary, which allows stronger and deeper social interaction between players, as they regularly interact with fellow guild members.

2.3.6 Comparing virtual environments, virtual reality and MMOs

Other virtual worlds are not games, but environments that may be used for communication and collaboration. According to Boulos, Hetherington, and Wheeler (2007) a virtual world is a computer-based, simulated multi-media environment, usually accessible via the Internet, designed so users can inhabit and interact via their own self-representations known as avatars. One such virtual environment is Second Life® (Linden Research, Inc., 2003).

Boulos et al. (2007) note that virtual worlds offer novel, intuitive ways to interact. These include navigating multi-media content; browsing information spaces and document collections in 3D virtual libraries; visiting new places and sampling new cultures; playing multi-player games; developing social and other skills; and attending and participating in live events.

Katz, Ring, Naveh, Kizony, Feintuch, and Weiss (2005) define Virtual Reality (VR) as a technology based on computerised simulation and real-time auditory, visual, and possibly haptic feedback. For Katz and team, the main advantage of using VR for rehabilitation is its inherent ecological validity, i.e. enabling the patient to experience realistic and meaningful therapy. Burdea (2003) notes that other advantages of using VR include
providing motivation, adapting to the patient, and ease of creating and retrieving data. Disadvantages include equipment costs and concerns over patient safety, although using the same hardware for multiple rehabilitation clients may mitigate the cost of VR hardware. The recent explosion of interest in VR, spearheaded by Oculus VR, has driven the development of consumer-level headsets, lowering the cost barrier (Avila and Bailey, 2014).

2.3.7 Therapeutic uses of MMOs and virtual environments

Fisher (1986) saw the potential of games to provide social interaction for disabled people, suggesting ‘perhaps the computer will be most effective when used socially, because so many brain damaged individuals feel isolated due to their disability’. Some research has investigated the rehabilitation potential for virtual environments, particularly Second Life; but surprisingly little research has been conducted into the therapeutic benefits of playing online multi-player games.

Boulos, Hetherington, and Wheeler (2007) state that ‘Second Life could be used to entertain older people and people with physical disabilities, and help them combat social isolation and loneliness’, but note that navigating a 3D environment may pose difficulties. This problem was solved by Hashimoto and team (Hashimoto, Ushiba, Kimura, Liu, and Tomita, 2010), who enabled a tetraplegic participant to walk around and chat in the Second Life virtual environment, showing that disabled people can have the same abilities as non-disabled people within virtual environments.

Galego and Simone (2007) developed a simple memory matching game played within Second Life, enabling the user and therapist to meet virtually. The game is controlled by the Nintendo Wii remote, and so exercises motor as well as cognitive skills. The authors state that playing this type of game can be a powerful rehabilitation activity, in areas such as upper extremity rehabilitation after a stroke.

According to Gaggioli, Gorini, and Riva (2007), using a virtual environment could enhance feelings of presence between therapist and patient, compared with other technologies like phone or email. These authors speculate that using avatars could create higher levels of interpersonal trust, and positively influence group cohesiveness. In the view of Wilkinson, Ang, and Goh (2008), virtual worlds offer the chance for therapists to meet clients in whatever environment is most appropriate. Gaggioli et al. (2007) state that multi-player online games are already used for therapeutic applications, giving the examples of an island within Second Life specifically designed for Asperger’s syndrome patients; and another area of Second Life intended for people affected by cerebral palsy.
Burdea (2003) reports that VR-based rehabilitation has been effective for sufferers of post-traumatic stress disorder. Treatment consists of exposure to virtual dangerous situations, such as helicopter flights over enemy territory, in order to desensitise the patient. Katz, Ring, Naveh, Kizony, Feintuch, and Weiss (2005) conducted a study into the effectiveness of a VR program to train stroke patients to cross roads safely. The participants suffered from unilateral spatial neglect as a result of the stroke. In this condition, patients will ignore objects or people on one side. The VR treatment was found to be effective in improving visual-spatial skills, and improved the ability to cross roads safely. More recently, the Oculus Rift VR headset has been successfully used as a distractor from pain for a child with severe burns (Hoffman et al., 2014).

2.3.8 Summary

Video games have been used therapeutically for many years in a wide variety of medical contexts. Virtual environments have also been used therapeutically, as non-threatening or convenient meeting places, allowing even severely disabled people to move and communicate. MMOs are at the intersection of video games and virtual environments, and, perhaps surprisingly, can bring out players’ pro-social, cooperative, altruistic natures.

2.4 Conclusions and questions

The prevalence of brain injury is increasing in the western world, with TBI described by one team as reaching ‘epidemic proportions’. Rehabilitation aims to restore as much independence as possible. The goal is to improve the person’s everyday functioning and participation by providing compensatory aids.

Video games have been found to provide therapeutic benefits in a variety of contexts, including serious games designed for TBI patients (Fisher, 1986) and stroke rehabilitation (e.g. Burke et al., 2009). Online multi-player games could be particularly beneficial to brain-injured people. Teamwork, cooperation and competition are all activities that are integral to multi-player games. Playing such games could improve a person's sense of belonging and friendship. Multi-player online games are a medium for communication between people as much as entertainment, with altruistic behaviour and friendships – even marriages – blossoming online. These social activities could be especially important to someone who feels isolation due to a disability.

Sohlberg and Mateer (2001) propose that improvements in cognitive and emotional abilities will proceed in tandem, not independently, following brain injury. In the general population, a link is seen between the emotionally damaging effects of loneliness and cognitive decline (Cacioppo and Hawkley, 2009). Video games have a long history of
being used therapeutically, with multi-player online video games fostering altruistic, pro-social behaviour. There is thus a link from the social, emotional therapeutic benefits of MMOs to counteracting the negative effects of isolation, to potential cognitive improvements.

This leads to the following research questions:

**Research Question 1:** Could playing a multi-player online video game provide a form of therapy for a brain-injured person? Would it improve cognitive function, emotional function, or both? Could any improvements be shown to be due to the social interactions fostered?

**Research Question 2:** If multi-player online video games can provide a cognitive therapeutic benefit, what are the essential features of these games that provide the benefit? Is it belonging to a community, playing cooperatively, behaving altruistically, etc.?

Research question 1 suggests several hypotheses. The first is that some cognitive skills can be measurably improved by exposure to a multi-player online game. We can envisage an experimental design to test this, with the hypothesis:

**H1:** Multiplayer online video games contribute to an improvement in cognitive function.

Research question 1 also asks if emotional function could be affected by such a game. This may not necessarily occur in tandem with any cognitive improvement, and so should be tested separately. Our definition of what we are measuring can be clarified to refer to loneliness and subjective satisfaction with life, two measures for which established instruments have been developed. This leads to two further hypotheses:

**H2:** Multiplayer online video games contribute to an improvement in subjective loneliness.

**H3:** Multiplayer online video games contribute to an improvement in subjective quality of life.

Should support for these hypotheses be found, the next stage would investigate the second research question, aiming to isolate the essential qualities of games that provide the benefits that have been proposed. One approach to this question is to add, remove and alter game features, and to measure the effect of these changes. A foundation for this approach is to use a game that is amenable to such changes as the research instrument.

The next chapter surveys research approaches that are suitable for investigating these questions and testing the hypotheses. Research methodologies are a key part of this review,
in order to find a suitable approach. Designing software for disabled people, software engineering methods, and game design principles are reviewed, with an emphasis on usability and accessibility.
3 Towards a Research Design

The last chapter identified research questions and hypotheses at the intersection of rehabilitation after brain injury, and the therapeutic potential of online multi-player video games. This chapter begins by reviewing research methods that could be used to experimentally test the hypotheses. Commonly used cognitive tests are surveyed, with the aim of finding those that could be used to measure any beneficial effects of multi-player online video games. The areas of video game design, human-computer interaction, designing for disabled people and software engineering methodologies are covered, to find suitable best practices that could be brought to bear.

3.1 Research methodologies

We begin by investigating research designs that could be employed in this research. Participant recruitment is of course essential to a clinical trial, and sample size may be a crucial success factor, depending on the research design employed. The single-subject research design, which is not highly dependent on the number of participants, is reviewed in detail.

3.1.1 Research designs

This section reviews established experimental designs: the retrospective study, the prospective study, and the randomised controlled trial. The case report is an in-depth description of an individual (Christensen, 2004). A retrospective (or case control) study is used to investigate factors that may prevent or cause disease. Patients are compared with a control group, with the aim of discovering factors that differ between the two groups. This type of study cannot establish causality because independent variables cannot be manipulated – the events have already happened. A prospective (or cohort) study compares a treatment group with a control group, which commonly receives a placebo treatment. A randomised controlled trial (RCT) is a prospective study in which the participants are randomly assigned to the treatment or control group. This randomisation counterbalances extraneous variables, and is described as the ‘gold standard’ for clinical trials (Babu, 2008).

3.1.2 Sample size

In a prospective study, the experimental design and associated plans for statistical analysis will imply a minimum number of participants. Greenfield (1996) notes that if the number of participants is too small to conclude anything useful, the study has exposed them to risk and inconvenience for no reason, and is thus unethical. Altman (1980) states
unequivocally: ‘If not enough subjects are available then the study should not be carried out or some additional source of subjects should be found’. Having too few participants is a common source of error in published studies: Moher, Dulberg, and Wells (1994) found that most RCTs with negative results suffered from sample sizes that were too small.

The power of a statistical test is the probability of rejecting a false null hypothesis (Christensen, 2004, p. 359), or, equally, that of correctly detecting a true difference between the control and experimental groups (Whitley and Ball, 2002). The power of a test depends on factors including the sample size; conversely, for a given power and required level of statistical significance, the required number of participants can be calculated. Altman (1980) declares that an underpowered study is ‘ethically indefensible’. Also, but more unusually, the sample size may be too high, wasting resources, and again exposing participants to risk and inconvenience unnecessarily (Whitley and Ball, 2002).

Sample size can be calculated precisely, but only if enough information is known. For example, to compare the mean in one group with the mean of another group, the sample size can be calculated given the required significance level and power, the smallest difference in means to be detected, and the standard deviation of the quantity being measured (Noordzij, Tripepi, Dekker, Zoccali, Tanck, and Jager, 2010). The greatest difficulty in calculating sample size is finding the value of the standard deviation without having already collected data. A value can be found by guessing, consulting an expert, obtaining a value from previous comparable studies, or carrying out a pilot study to calculate the value (Kumar, 1999). Another way of arriving at sample size is to use a nomogram, such as that devised by Altman (1980). Using this graphical device, we draw a line between the required statistical power and the standardised difference between values in two treatment groups. We can then read off the number of participants required for the required significance value.

### 3.1.3 Participant recruitment

From the above, sample size is critical for the validity of an RCT. The method of recruiting participants can also affect the result. Brase, Fiddick, and Harries (2006) found wide variation in the results of a psychological test that has been replicated multiple times. This was demonstrated to be partly due to the form of the question posed, but also to the sampling methodology: the general level of academic attainment of the participants was found to be a factor; so too was whether or not the participants had been paid for their involvement. Brase et al. conclude that relative performance within a group is more reliable than absolute scores.
Blanton et al. (2006) found that ‘participant recruitment is considered the most difficult aspect of the research process’. Those authors describe the difficulties of recruitment, and detail the efforts made to recruit for the Extremity Constraint-Induced Therapy Evaluation (EXCITE) trial, a 5-year rehabilitation research project conducted across six sites in the US. The study had 222 participants, out of a goal of 240. Blanton and team identified a number of tools and strategies as part of their recruitment plan. A website for the study was developed, containing information for potential participants. A brochure was distributed, along with marketing gifts branded with the study logo (T-shirts, pens, mugs, etc.) The most effective recruitment strategy was found to be presentations at physiotherapy and occupational therapy clinics, with follow-up phone calls. The authors state that the presentations were so successful because group members frequently ‘spread the word’, passing on news of the study to other potential participants. Blanton et al. (2006) also considered retention, and report that several of the sites provided participants with notebooks and a personalised calendar, a copy of the signed consent form, a letter of gratitude, etc. Most importantly though, Blanton et al. (2006) suggest that retention may best be improved by building personal relationships with the participants, and maintaining communication.

3.1.4 Single-subject research design, SSRD

Given the difficulty of recruitment noted in the previous section, it may be that the number of participants in the current study is low. This section looks at a research design that requires only the absolute minimum number of participants, and so this may be the most applicable design.

A single-subject research design (SSRD), also called ‘single-case’ or ‘n=1’ design, requires only a single participant. An SSRD relies on repeated measurements of the dependent variable (Christensen, 2004). The measurements taken before the introduction of the treatment condition are called the baseline. An ideal baseline will have no trend and little variability. The second phase introduces the treatment condition. The third phase is to withdraw the treatment condition, giving an ‘A-B-A’ design. This step is considered crucial in an SSRD to demonstrate that the treatment condition caused the dependent variable change, refuting rival hypotheses. The fourth step is to reinstate the treatment (an ‘A-B-A-B’ design).

The problem with the A-B-A design is that withdrawing treatment may not cause a reversal to the baseline if the result of the treatment is permanent. The multiple-baseline variant may be used in SSRDs where the treatment is non-reversible, because this design does not rely on withdrawal of the treatment condition (Christensen, 2004). Multiple
baseline data is collected, either for the same behaviour for multiple individuals, or
different behaviours for the same individual. The experimental treatment is administered at
successive intervals. Corresponding changes in the dependent variable would make rival
hypotheses increasingly unlikely as more treatments are administered. The reversal and
multiple-baseline designs provide the strongest evidence of causal relationships that can be
attained from single-subject designs (Christensen, 2004).

An SSRD is appropriate when research funds are scarce; when research questions address
the process of change; and research questions aim to find out if a treatment would work for
a particular patient (Janosky, 2005). The major limitation of a single-case design is that the
results cannot be generalised. A multiple-baseline design can strengthen external validity.
Janosky (2005) suggests further ways in which external validity can be increased: by
choosing a participant who is representative of the type of patient who would undergo the
treatment; and by replicating the study with variation in researcher, participant, or
practices.

### 3.1.5 Statistical analysis of SSRDs

In an SSRD, there is an expectation that results collected in the different phases (i.e.,
intervention or non-intervention) will be visually distinct when plotted. Discussing
possible statistical analyses of SSRD data, Houle (2009) states ‘There is simply no
replacement for the information provided by graphing the outcome variable as it varies
over time’. Christensen (2004) notes that statistical analysis will add little to the
interpretation of the results if a stable baseline and limited variability can be achieved, but
if the results are not obvious by visual inspection, statistical analysis would be of some use.

So, although visual inspection is a standard and irreplaceable technique for analysing the
results of an SSRD (Houle, 2009), use of statistical tests is desirable as it may add an extra
dimension to the analysis. Unfortunately data collected in an SSRD may violate the
assumptions of a statistical test. This arises due to the non-independent, non-normally
distributed nature of time-series data. Some statistical techniques do not apply in the case
of time-series data, where contiguous values are auto-correlated.

### 3.1.6 Proposed experimental design

To summarise the above, the clinical research method used in a study depends on factors
including availability of data, the likely number of participants, and ethical considerations.
This means that the ‘gold standard’ of an RCT is not necessarily the right choice. Single-
subject designs are suitable for studies with low participant numbers, because each
participant is effectively their own control. The external validity of a study employing this method is lower, and inferential statistics are less likely to be applicable.

At this stage, an experimental approach is proposed to measure the effect of playing a multi-player online video game. A SSRD would be used, if the number of participants were low. If enough participants were recruited, other designs would be considered also. The independent variable would be whether or not the participant plays a multi-player online video game. This would vary according to the phase of the experiment, in an A-B-A or A-B-A-B pattern. The next step is to identify dependent variables. It is proposed here that emotional wellbeing, and feelings of loneliness in particular, should improve due to the social interactions present in a multi-player game. Cognitive improvement would be seen alongside the emotional, if it is indeed the case that emotional and cognitive recovery are inseparable, as proposed by Sohlberg and Mateer (2001). The dependent variables should measure each participant’s sense of wellbeing, and their cognitive skills.

3.2 Cognitive tests

Given the proposed experimental design, this section identifies tests that participants could take, that could measure the effects on cognitive ability of playing a multi-player online game. These tests would be the dependent variables in the proposed experiment. An ideal test would have these characteristics (Shulman, 2000):

1. Quick to administer;
2. Well-tolerated and acceptable to patients;
3. Easy to score;
4. Relatively independent of culture, language and education;
5. Have good inter-rater and test-retest reliability;
6. Have high levels of sensitivity and specificity;
7. Have concurrent validity (correlation with measures of severity and other rating scores); and
8. Have predictive validity.

Cullen et al. (2007) undertook a review of screening tests for cognitive impairment. Their focus was on dementia (Alzheimer’s disease), aiming to identify cognitive screening tests, and assess their suitability for three purposes: brief assessment in a doctor's office, large-scale screening, and domain-specific screening to guide further assessment. They identified six core cognitive domains that should be covered by any comprehensive screening instrument:

1. Attention/working memory,
2. New verbal learning and recall,
3. Expressive language,
4. Visual construction,
5. Executive function, and
6. Abstract reasoning.

Further criteria are introduced here, specific to the current research. Participant recruitment is viewed as a difficult part of the research process, which can be exacerbated further by location. Blanton et al. (2006) found that ‘transportation emerged as an important factor and the most critical (non-medically related) recruitment concern, as participants were required to travel to and from the research site on a frequent basis.’ If the current experimental design required the researcher to be physically present to administer tests, then the pool of participants would be reduced. To avoid this, it is proposed that the tests be administered by software, on the same computer or device as the multi-player video game, removing location and mobility as barriers to participation. This design means the criteria for suitable tests are narrowed: for a test to be suitable, it must be possible to implement it in software, and for the participant to take the test with no researcher present.

The viability of this plan is supported by the availability of suites of computerised cognitive tests, which could be administered remotely. An example of a commercial offering is the Cantab test suite, developed by Cambridge Cognition, Ltd. (Cambridge Cognition, Ltd, 2015). PEBL, the psychology experiment building language, is a free, open-source tool for building cognitive tests, and has a catalogue of tests that have already been created (Mueller and Piper, 2014). The existence of these computerised test suites indicates that it is viable to use computerised tests in the present research. The next sections review cognitive tests, which are commonly used in clinical practice, and may be suitable as tests in the proposed experimental design.

3.2.1 The Mini-Mental State Examination (MMSE)

Folstein, Folstein and McHugh (1975) developed the MMSE as an instrument that would be quicker to administer than existing cognitive tests of the day (hence the ‘mini’). The MMSE consists of eleven items, with a maximum mark of 30. It takes 5 to 10 minutes to administer, but is not timed. It concentrates only on cognitive functions, but in that domain is described as ‘thorough’ by its authors. There are two sections. The first covers orientation, memory, and attention, requiring verbal responses. The second part involves reading, writing and drawing. Overall, the MMSE is validated as a test of cognitive function, which distinguishes between people with and without cognitive disturbance, and is sensitive to changes as patients recover.
The eleven MMSE items are as follows. As can be seen, some variation and flexibility in the questions is allowed.

1. What is the year/season/date/day/month?
2. Where are we: state/county/town/hospital/floor?
3. Name 3 objects: 1 second to say each, then ask patient all 3.
4. Serial 7s (see below), or spell ‘world’ backwards.
5. Ask for the 3 objects repeated above.
6. Name a pencil, and a watch.
7. Repeat: ‘no ifs, ands or buts’.
8. Follow a 3-stage command: ‘take a paper in your right hand, fold it, and put it on the floor’.
9. Read and obey the following: ‘CLOSE YOUR EYES’.
10. Write a sentence.
11. Copy a design of two intersecting pentagons.

Serial 7s (item 4) refers to the task of counting backwards from 100 in multiples of 7. Task 11 asks the subject to copy a figure consisting of two overlapping pentagons, the overlapping region forming a rhombus. This test assesses motor ability (Fountoulakis et al., 2011) and is sensitive to visual-constructional apraxia (Mitolo, Salmon, Gardini, Galasko, Grossi, and Caffarra, 2014). Figure 3.1 (overleaf) shows the figure to copy, Figure 3.2 (overleaf) shows examples of acceptable attempts, scoring 1 point, and Figure 3.3 (overleaf) shows some unacceptable attempts: the point is given only if two five-sided figures are drawn, overlapping to form a four-sided embedded figure (Ridha and Rossor, 2005).

The MMSE is the most commonly used cognitive test in clinical practice (Cordell et al., 2013; Bossers, van der Woude, Boersma, Scherder, and van Heuvelen, 2012; Cullen, O’Neill, Evans, Coen, and Lawlor, 2007; Kipps and Hodges, 2005). Unfortunately, the MMSE is not free to use or reproduce. Newman and Feldman (2011) chart the change in status of the MMSE, from once being freely available to now being vigorously protected under copyright law. Fong et al. (2011) developed the ‘Sweet 16’ (named for its maximum score of 16) as an improved, free, alternative to the MMSE. According to Newman and Feldman, the Sweet 16 was ‘removed from the Internet’ due to a ‘takedown’ by the copyright holders of the MMSE.
Figure 3.1. Pentagon copying in the Mini Mental State Examination: the original figure (Ridha and Rossor, 2005).

Figure 3.2. Pentagon copying in the Mini Mental State Examination: acceptable copies (Ridha and Rossor, 2005).

Figure 3.3. Pentagon copying in the Mini Mental State Examination: unacceptable copies (Ridha and Rossor, 2005).

From the point of view of the present research, using the MMSE would present another problem. As can be seen from the MMSE items, this test is intended to be administered by a clinician, in person. The test does not lend itself well to a software implementation.

After the MMSE, the most common assessment tools are the clock-drawing test, delayed-word recall, the verbal fluency test, the similarities test, and the trail-making test (Shulman et al., 2006). In a 2012 review of tests used for cognitive assessment of dementia patients, Bossers et al. found that the Stroop colour-word interference test was used in two studies, compared to 54 studies for MMSE; a reaction time test was used in one study, and a digit
cancellation task was used in one study (Bossers, van der Woude, Boersma, Scherder, and van Heuvelen, 2012).

### 3.2.2 Clock-drawing test (CDT)

Although there exist variations in instructions and scoring systems, a common form of the clock-drawing test (CDT) is to present the subject with a pre-drawn circle of about 10 cm diameter. The instructions given are: ‘this circle represents a clock face. Please put in the numbers so it looks like a clock and then set the time to 10 minutes past 11’. This time is considered useful because it includes both visual fields, and requires inhibition of the ‘pull’ towards the number 10 (Shulman, 2000). Shulman speculates that the CDT could be the ideal cognitive test, finding that it has high correlation with the MMSE, and is sensitive to cognitive change, with good predictive validity. The clock-drawing task taps into a wide range of cognitive abilities, and is quick and easy to administer. He writes ‘the clock-drawing test is complementary to the widely used and validated Mini Mental State Examination and should provide a significant advance in the early detection of dementia and in monitoring cognitive change’. Figure 3.4 (overleaf) shows examples of clocks drawn by patients, and the severity score awarded by a clinician, from 0 (most severe brain injury) to 5 (least severe). This is one of five CDT scoring systems, with the others rating clocks on scales of 0-2, 0-1, 1-10, and 1-6 (Kørner, Lauritzen, Nilsson, Lolk, and Christensen, 2012). In the case of the 0-5 scale, the scoring criteria are as shown in Table 3.1 (overleaf).

There are two problems with implementing the clock-drawing test in software. The first is that a standard PC does not have a suitable input device, although a touch screen or tablet with stylus could be used. The second, much harder problem is how the software would rate the clock drawn by the participant, a task requiring clinical judgement.
Figure 3.4. Clock drawing examples for each of the severity scores from 0 to 5 (Shulman, 2000).

Table 3.1. Scoring manual for the clock-drawing test on 0-5 scale (Kørner, Lauritzen, Nilsson, Lolk, and Christensen, 2012).

<table>
<thead>
<tr>
<th>Score</th>
<th>Benchmark</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>Numbers and hands are correctly placed</td>
</tr>
<tr>
<td>4</td>
<td>Mild visuo-spatial errors</td>
</tr>
<tr>
<td>3</td>
<td>Clear-cut errors in time given</td>
</tr>
<tr>
<td>2</td>
<td>Moderate visuo-spatial errors</td>
</tr>
<tr>
<td>1</td>
<td>Marked visuo-spatial errors</td>
</tr>
<tr>
<td>0</td>
<td>Neither numbers nor hands are remotely correctly placed</td>
</tr>
</tbody>
</table>
3.2.3 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 and Riccio, 2004) - the ability to suppress a learned response in favour of an unusual response.

The Stroop tests are comprised of three tasks: the colour test, word test, and colour-word test. In the colour test, the participant is asked to identify the colours of rectangles. In the word test, the participant is asked to name the colour of a number of words, where each word is printed in the colour that it names. Figure 3.5 shows an example of the Stroop word test. In the colour-word test, the participant should name the colour of the word, not the word itself. An example of this test is shown in Figure 3.6. The difficulty of this task was discovered by Stroop (1935), and is ascribed to interference between the two separate cognitive tasks of identifying a colour and reading a word.

<table>
<thead>
<tr>
<th>RED</th>
<th>GREEN</th>
<th>BLUE</th>
<th>YELLOW</th>
<th>PINK</th>
</tr>
</thead>
<tbody>
<tr>
<td>ORANGE</td>
<td>BLUE</td>
<td>GREEN</td>
<td>BLUE</td>
<td>WHITE</td>
</tr>
<tr>
<td>GREEN</td>
<td>YELLOW</td>
<td>ORANGE</td>
<td>BLUE</td>
<td>WHITE</td>
</tr>
<tr>
<td>BROWN</td>
<td>RED</td>
<td>BLUE</td>
<td>YELLOW</td>
<td>GREEN</td>
</tr>
<tr>
<td>PINK</td>
<td>YELLOW</td>
<td>GREEN</td>
<td>BLUE</td>
<td>RED</td>
</tr>
</tbody>
</table>

Figure 3.5. Stroop word test (Chudler, 2015).

<table>
<thead>
<tr>
<th>RED</th>
<th>GREEN</th>
<th>BLUE</th>
<th>YELLOW</th>
<th>PINK</th>
</tr>
</thead>
<tbody>
<tr>
<td>ORANGE</td>
<td>BLUE</td>
<td>GREEN</td>
<td>BLUE</td>
<td>WHITE</td>
</tr>
<tr>
<td>GREEN</td>
<td>YELLOW</td>
<td>ORANGE</td>
<td>BLUE</td>
<td>WHITE</td>
</tr>
<tr>
<td>BROWN</td>
<td>RED</td>
<td>BLUE</td>
<td>YELLOW</td>
<td>GREEN</td>
</tr>
<tr>
<td>PINK</td>
<td>YELLOW</td>
<td>GREEN</td>
<td>BLUE</td>
<td>RED</td>
</tr>
</tbody>
</table>

Figure 3.6. Stroop colour-word test: the task is to name the colour of each word, not to read the word (Chudler, 2015).
Homack and Riccio (2004) point out that there are numerous variants of the Stroop tests and no ‘official’ version. The validity of using computerised Stroop tests is supported by the findings of Hepp, Maier, Hermle, and Spitzer (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 a researcher recorded responses manually. 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 or pointing version could be more accurate than oral responses: Hepp et al. (1996) recommend that computerised versions of the Stroop task should be used.

3.2.4 Trail-making tests

Trail-making tests are another common assessment tool (Shulman et al., 2006), validated as an indicator of brain injury (Reitan, 1958) and believed to reflect a wide range of cognitive abilities:

‘Attention, visual search and scanning, sequencing and shifting, psychomotor speed, abstraction, flexibility, ability to execute and modify a plan of action, and ability to maintain two trains of thought simultaneously’ (Salthouse, 2011).

In these tests, the participant uses a pencil to connect randomly distributed numbered or lettered circles in order. In the simpler 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’ (Zakzanis, Mraz, and Graham, 2005). Figure 3.7 (overleaf) shows example trail-making tests, and Figure 3.8 (overleaf) shows the trail-making test implemented as software, in PEBL (Mueller, 2012).

The ‘Connections’ variant of the trail-making test, developed by Salthouse et al. (2000), attempted to minimise the influence of visual search and hand movement, the better to focus on the other cognitive processes believed to be involved in the test. This was accomplished by positioning successive targets adjacently. Salthouse (2011) found that the predominant cognitive capabilities measured by the trail-making tasks were speed of thought, reasoning, and spatial visualisation.
Figure 3.7. Trail-making tests A and B (Winner, 2009).

Figure 3.8. Trail making test implemented in PEBL (Mueller, 2012).
3.2.5 Letter cancellation tests

In a letter cancellation test, the participant is required to cross out the occurrences of one particular letter, randomly distributed in a grid of other random letters. Rorden and Karnath (2010) state that cancellation tests are used to diagnose spatial neglect, and that the popularity of these tests may be due to the ease of description and speed of administration. Malec et al. (1984) describe their letter and symbol cancellation tests as having a grid of 6 rows of 52 random letters. The task was to cross out 32 randomly distributed M letters, and 36 random Σ symbols. Figure 3.9 shows an example of a letter cancellation test sheet. It can be seen that cancellation tests could be implemented in software, but, using a standard mouse, performance would depend greatly on motor coordination.

![Letter cancellation test sheet](image)

Figure 3.9. An example of a letter cancellation test sheet (Canadian Partnership for Stroke Recovery, Letter cancellation test sheet).

3.3 Tests to measure quality of life and loneliness

This research proposes that any improvements in cognitive skills would be matched by improvements in the loneliness felt by a player of multi-player online games. Weiss (1973) identified two types of loneliness: emotional and social. Emotional isolation is the absence of close emotional attachment, while social isolation is the absence of socially integrated relationships. The symptoms of emotional loneliness are anxiety and emptiness; the
symptoms of social isolation are feelings of boredom, aimlessness, and marginality. From these definitions, it is proposed here that of the two types of loneliness, social loneliness is more likely to be affected by the activities of a multi-player online game. For a brain-injured player, a positive outcome would be an improvement in perceived quality of life. To pursue this line of research further, this section looks at research instruments that could be used to measure the degree of loneliness, and happiness, felt by a person. They vary depending on the underlying conceptual model, and also how burdensome they are for the respondent to complete.

3.3.1 UCLA Loneliness Scale
Russell, Peplau and Ferguson (1978) developed the UCLA Loneliness Scale. This scale consists of twenty statements, each of which can be answered on a four-valued scale: ‘never’, ‘rarely’, ‘sometimes’, and ‘always’. This scale is frequently used as a loneliness measure (Hawkley and Cacioppo, 2010; Chlipala, 2008), despite being regarded as lengthy.

Shorter instruments have been developed, including shortened versions of the UCLA scale (Chlipala, 2008). Pinquart and Sörensen (2001) found that the most frequent measure of loneliness is a single item. Results from this type of instrument may suffer from under-reporting, as the question may be too explicit, given that loneliness has negative connotations. Other loneliness scales include the NYU Loneliness Scale. This scale has 8 items, with the response to each varying from a four- to seven-point scale. It is infrequently used (Chlipala, 2008).

3.3.2 De Jong Gierveld 6 item Loneliness Scale
Another measure, the De Jong Gierveld Loneliness Scale, was developed by De Jong Gierveld and Kamphuis (1985). The original scale has 34 items. A shortened version consisting of only 6 items was later developed (De Jong Gierveld and Van Tilburg, 2006). This scale attempts to measure the social and emotional components of loneliness, as identified by Weiss (1973). In a meta-analysis of 149 loneliness studies published between 1948 and 1999, Pinquart and Sörensen (2001) found that the De Jong Gierveld scale was used less frequently than the UCLA scale, but more frequently than others.

The 6-item test consists of 6 questions measured on a five-point scale. The questions are listed in Table 3.2 (overleaf).
Table 3.2. De Jong Gierveld 6-point Loneliness Scale questions.

<table>
<thead>
<tr>
<th>Question</th>
<th>Text</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>I experience a general sense of emptiness.</td>
</tr>
<tr>
<td>2</td>
<td>I miss having people around.</td>
</tr>
<tr>
<td>3</td>
<td>I often feel rejected.</td>
</tr>
<tr>
<td>4</td>
<td>There are plenty of people I can rely on when I have problems.</td>
</tr>
<tr>
<td>5</td>
<td>There are many people I can trust completely.</td>
</tr>
<tr>
<td>6</td>
<td>There are enough people I feel close to.</td>
</tr>
</tbody>
</table>

The response to each question is one of the following:

- Yes!
- Yes
- More or less
- No
- No!

De Jong Gierveld and Van Tilberg (2006) explain how the results are processed. The neutral and positive answers are counted for questions 1, 2, and 3. This is the emotional loneliness score. This score is only valid if all of these questions are answered. The neutral and negative answers are counted for questions 4, 5, and 6. This is the social loneliness score. This score is only valid if all of these questions are answered. The total loneliness score is the sum of the emotional loneliness score and social loneliness score. This is only valid if the total number of unanswered questions is zero or one.

### 3.3.3 Satisfaction With Life Scale

The Satisfaction With Life Scale was developed by Diener, Emmons, Larsen, and Griffin (1985) and has been used and evaluated many times since then (Pavot and Diener, 2008). It is another short instrument (5 items). Participants answer five questions on a 7-point scale. The questions are listed in Table 3.3 (overleaf).
Table 3.3. Satisfaction With Life Scale questions.

<table>
<thead>
<tr>
<th>Question</th>
<th>Text</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>In most ways my life is close to my ideal.</td>
</tr>
<tr>
<td>2</td>
<td>The conditions of my life are excellent.</td>
</tr>
<tr>
<td>3</td>
<td>I am satisfied with my life.</td>
</tr>
<tr>
<td>4</td>
<td>So far I have gotten the important things I want in life.</td>
</tr>
<tr>
<td>5</td>
<td>If I could live my life over, I would change almost nothing.</td>
</tr>
</tbody>
</table>

The response to each question is one of the following:

- Strongly disagree,
- Disagree,
- Slightly disagree,
- Neither agree nor disagree,
- Slightly agree,
- Agree,
- Strongly agree.

A numeric value is assigned to the response to each question, from 1 (‘Strongly disagree’) to 7 (‘Strongly agree’), giving an overall score, which is the sum. This total score ranges from 5 (lowest satisfaction) to 35 (highest satisfaction). Kobau, Sniezek, Zack, Lucas, and Burns (2010) validated a slightly simplified version of this test, in which the seven responses are reduced to five, lowering the cognitive burden for the respondents. This results in an overall score that can range from 5 to 25.

We have reviewed research designs, and cognitive tests which could be used in an experiment to test whether or not multi-player online video games can have a beneficial effect. In the current research, it is anticipated that game software will be developed, or modified. The next sections look at principles and methods of human-computer interaction, software development, and video game design. These will guide the development of game software required as a research instrument.
3.4 Human-Computer Interaction: principles and methodologies

Usability is a key requirement for games as much as any other software. This section reviews the subject of human-computer interaction (HCI), so that the principles described here can be applied to the game software.

3.4.1 Usability Engineering

Nielsen (1993) defines usability as consisting of five attributes. The system should be:

- Easy to learn,
- Productive to use,
- Easy to memorise,
- Result in a low error rate, and
- Provide satisfaction to the user.

These traits are measurable, allowing interfaces to be assessed and compared. Usability Engineering is a methodology for building usable interfaces (Nielsen, 1993), the major feature of which is the definition of explicit usability criteria early on in the design process, which are used to judge the final system (Dix, Finlay, Abowd, and Beale, 2004).

Usability testing is fundamental to Usability Engineering, providing direct observational data. Possibly the most valuable technique in usability testing is ‘thinking aloud’ (Nielsen, 1993, p. 195). This is a traditional psychological research method where participants verbalise their thoughts as they perform tasks. Thinking aloud provides rich qualitative data from even a small number of participants, but is not without disadvantages, as the users may be distracted from properly performing their task by having to continually verbalise their thoughts.

Other methods of getting user data are ‘fly on the wall’ observation, questionnaires, interviews, focus groups, and logging actual use. Questionnaires, interviews and focus groups are indirect methods, in that they do not study the interface itself, but the opinions of the user about the interface. This is appropriate in the case of user satisfaction, but for other criteria, data should take precedence over opinion. Logging data from real users in the field is recommended (Nielsen, 1993), to gather statistics such as how many features are used, or the rate of errors; to allow playback of a user’s actions to analyse usage, and to find usability problems that are not apparent during observations. A high frequency of a particular kind of error would indicate an area that could be improved in a future release.
3.4.2 Contextual enquiry and design

Contextual Design is another user interface design methodology (Holtzblatt, Wendell, and Wood, 2005). The first step is Contextual Inquiry, consisting of field interviews with prospective users. The purpose is to discover real world practice and activities, rather than an idealised or official version, and so consists of observation and discussion of the user’s tasks.

Subsequent steps lead to the development of ‘Personas’, descriptions of a fictional, typical user, created by compositing the data drawn from the contextual inquiries. Storyboarding is used to document how users will accomplish their tasks using the redesigned system.

The new interface is first constructed as a paper prototype, made of Post-It® Notes or similar. This is used to get feedback on the proposed design. Paper Prototype Interviews with users are used to find flaws in the design, leading to the next iteration. Paper prototypes are also known as ‘lo-fi’ prototypes, while designs implemented on a computer are known as ‘hi-fi’ prototypes (Snyder, 2003).

3.4.3 Iterative prototyping

Dix et al. (2004) note that iterative prototyping is the universally accepted approach to interaction design, because the only way to be sure about the usability of an interface is to build it and test with real users. The design is then modified to correct false assumptions, and the process is repeated. Prototypes simulate some features of the final intended system.

There are three main approaches to prototyping: throw-away, incremental, and evolutionary. Throw-away prototypes are discarded and rewritten once they have been used to determine usability flaws. In incremental prototyping, the final product is developed as separate smaller components, which are added one at a time. Evolutionary prototyping builds more functionality into the prototype at each stage.

One pitfall of iteration is that it is a hill climbing exercise: the process could become stuck at a local rather than global maximum. A good starting point is needed, so starting with several design ideas, rather than one, is recommended (Dix et al., 2004).
3.5 Designing for disabled people

All the HCI approaches described in the previous section are based on non-disabled users, with normal ranges of mobility, sensory and cognitive function. This section looks at considerations arising when designing for people with disabilities.

Newell (1995) takes the view that disability is a function of a person’s environment, as, for example, noise or poor light will result in the equivalent of deafness and visual impairment. Similarly, other inhospitable environments could cause effects such as mobility impairment and tactile insensitivity, and fatigue and stress will cause reductions in cognitive abilities.

An important principle of HCI design is to recognise that there is a spectrum of variability among potential users. Users of a system will range from novice to expert in their skills for a particular task. Much HCI work has often appeared to assume that all users are the same except for variation along the novice-expert axis (Newell, 1995). Unfortunately, designing for exaggerated differences between disabled and non-disabled people can also lead to design faults. Design engineers must study the important characteristics of their users. Newell (1995) advises that ‘very detailed’ consideration of the users’ needs and wants is required.

3.5.1 Universal design

Universal design is defined by Dix et al. as ‘the process of designing products so that they can be used by as many people as possible in as many situations as possible’ (2004, p. 366). This can be achieved by designing-in redundancy, or by supporting assistive technologies. Universal design is based on seven principles proposed by Story, Mueller, and Mace (1998). These are:

1. Equitability – no user should be excluded or stigmatised,
2. Flexibility,
3. Simplicity,
4. Clarity,
5. Error tolerance,
6. Low physical effort requirements, and
7. Sufficient size and space for use.

Multi-modal interaction is recommended, providing access to information through more than one mode of interaction. This both enhances the richness of interaction, and also guards against excluding a group of users.
A good example of the success of universal design is the addition to web sites of features for blind and low-vision users. Adding font-size customisability and screen reader tags to a web site is standard practice (Barlet and Spohn, 2012).

3.5.2 Working with brain-injured participants

Working with brain-injured participants presents particular methodological challenges. Malec et al. (1984) note that standard experimental designs are not often appropriate when comparing the abilities of this group of people, as recovery rates can be uneven and unpredictable. Gnanayutham and George (2009) report difficulties that arise when working with brain-injured participants due to individual abilities and disabilities, the effects of medication, attention span, and frustration.

Gnanayutham and George (2009) propose an approach to developing user interfaces for neurorehabilitation software, attempting to cater for the largest group of users, reducing the scope until a feasible solution is found:

- The first step is to determine if a universal interface is feasible.
- If not, determine if interfaces for groups of people or classes of injury can be developed.
- If this again is not possible, determine if personalised interfaces can be developed.
- At this level, the personalised interface may be a customised version of the same paradigm, or a different paradigm for each person.

This approach would seek to develop an interface that benefits the entire user population, if feasible, and then move to smaller subsets of the population if required. The benefit of such an approach is that a universal interface would benefit the largest number of people for the lowest development cost. A potential disadvantage could be that a ‘one size fits all’ interface effectively targets the lowest common denominator. This pitfall may be avoided by making the interface adaptable.

3.5.3 Assistive technologies

An Assistive Technology (AT) is any technology designed to help provide independence for disabled people. The World Health Organization defines AT as ‘an umbrella term for any device or system that allows individuals to perform tasks they would otherwise be unable to do or increases the ease and safety with which tasks can be performed’ (World Health Organization, 2004). This broad definition encompasses a wide variety of technology types, such as sitting, standing and ambulatory aids, wheelchairs and ramps, prostheses, and hearing aids, as well as hi-tech computer interfaces. The single most
widely used AT device is probably the walking stick (Russell, Hendershot, LeClere, Howie, and Adler, 1997).

Augmented and alternative communication (AAC) systems help disabled people communicate. These range from lo-tech solutions, such as a picture board, to hi-tech direct selection or scanning systems (Glennen, 1997). Direct selection systems include keyboards; switch-activated systems; optical pointers; and eye gaze systems. Scanning systems can be used by people with more limited physical abilities, or when direct selection would cause fatigue. The limitation is data rate, so direct selection techniques are usually preferred.

3.6 Software Engineering methodologies

In this section, software engineering methodologies are reviewed, which may be applied to software development tasks that arise in this research. It was the intention of the researcher to enlist volunteers who could help with the software development. Even a small team can benefit from some degree of process: McConnell (1998) advises that it is safer to err on the side of too much process, and later relax it, rather than try, too late, to apply more process to a failing project. The smallest possible software development team is, of course, one person. Even in this case, some kind of process is necessary, to ensure that the software is completed, to the required standard, by the given deadline.

The first software engineering conference was held in 1967, organised by the NATO Science Committee. At the time, the term was intentionally contentious, implying that the process of creating software required more discipline (Jackson, 1995, p. 188).

Somerville (2011) defines software engineering as ‘an engineering discipline concerned with all aspects of software production’. According to Jackson (1995, p. 58), ‘the central activity of software development is description’. McConnell (1998, p. 31) regards software development as gradual reduction in uncertainty, while Brooks (1982) memorably likened large-scale software engineering to a tar pit. The purpose of any methodology is to impose a disciplined process on software development, aiming to make it more predictable and efficient (Meso and Jain, 2006).

3.6.1 Problems with software development

Abrahamsson, Salo, Ronkainen, and Warsta (2002) note that many software development methodologies have been proposed over the years, few of which have survived. This may lead practicing programmers to regard new approaches with some scepticism, leading organisations to not follow a specific software engineering process, but rather rely on improvised methods. A follow-on effect of this is that it is hard to improve subsequent
projects. McConnell (1998) noted that some software developers view the very notion of ‘process’ as a restriction on their creativity.

Unfortunately, following a methodology does not guarantee success. Jones (2004) analysed 250 large software projects undertaken between 1995 and 2004. Only 10% of the projects were considered successful in terms of schedule, cost and quality. About 70% of the projects had major schedule overrun or were terminated before completion, despite following well-known methodologies.

Mistakes at the programming level can cause disasters. The European Space Agency’s Ariane 5 rocket exploded after take-off in 1996, due to a programming error (Bryant and O’Hallaron, 2003, p. 97). Errors in the software of the Therac-25, a radiation therapy device, caused six known deaths by administering massive radiation overdoses (Leveson and Turner, 1993). These examples are clearly catastrophic failures, but for any given piece of software, some worst-case behaviour can be envisaged.

### 3.6.2 Software development process models: from ‘heavyweight’ to ‘lightweight’

This section looks at software development processes which might be usable in developing software for the present research project. Sommerville (2011) notes that there is no ideal software development process because of the diversity of software processes and requirements, and the need for judgement and creativity. However, fundamental activities are common to all software processes: specification, design and implementation, validation, and evolution. Software development methodologies have shifted over time, from earlier more ‘heavyweight’ approaches to more recent ‘lightweight’ or ‘agile’ processes.

In the waterfall process model (Royce, 1970), software is developed in a strict sequence of activities. These activities are: requirements definition, system and software design, implementation and unit testing, integration and system testing, and operation and maintenance. This model has been criticised for its inflexibility (Sommerville, 2011, p. 31; Beck and Andres, 2005, p. 87) and for leaving serious risks unaddressed until the integration and system test phase (Jacobson, Booch, and Rumbaugh, 1999). Somerville (2011) asserts that the waterfall model is used for larger projects where the requirements are well understood.

Addressing the issues of the waterfall model, Boehm (1988) proposed a spiral model for software development. In this model the distance from the centre of the spiral represents the cumulative cost of the project. One circuit comprises four major phases: determination of objectives, evaluation of alternatives, and identifying and resolving risks, development
and verification; and planning the next phases. Boehm (1988) states that the defining feature of the spiral model is that it is ‘risk-driven’ rather than driven by documentation or code. This risk-driven approach may avoid the problems of other models, but exposes the project to a reliance on risk-assessment expertise.

McConnell (1998) notes the success of the NASA Software Engineering Lab (SEL) in developing high-quality software (Basili, McGarry, Pajerski, and Zelkowitz, 2002), and advocates a staged release approach. This borrows from the SEL Recommended Approach to Software Development (Landis et al., 1992), and has similarities with the waterfall approach, first requiring requirements development and an architectural design. The process then becomes iterative, where each iteration brings the product to a releasable state, with a subset of the final feature set implemented.

Since the mid 1990s, a succession of lightweight, ‘agile’ software development processes have been developed. The ‘agile manifesto’ declares that it is preferable to focus on people rather than processes, code over documentation, collaboration over contracts, and responsiveness to change over planning (Beck et al., 2001).

Extreme Programming (‘XP’) is described by Beck and Andres (2005) as a lightweight methodology that assumes requirements will be vague and changing. XP is described as encompassing:

- A philosophy of software development based on values of communication, feedback, simplicity, courage and respect,
- Principles and best practices, and
- A community that shares the XP ‘values’ and many of the same practices.

XP also introduced the then novel practice of pair programming, where two programmers work together on the same piece of code.

Scrum is an agile methodology focusing on team management, rather than programming techniques, developed by Sutherland and Schwaber (2011) and first used in 1993. A Scrum-based project has three phases, called ‘pre-game’, ‘game’ or ‘development’, and ‘post-game’. The development phase consists of a series of ‘sprints’, where functionality is developed or enhanced. One sprint lasts from one week to one month. XP and Scrum are the most widely used of the agile methods, with Scrum used in over 75% of agile projects worldwide (Sutherland and Schwaber, 2011).

Other agile methodologies have been proposed, but have not enjoyed the same level of acceptance as Scrum. These include Crystal (Cockburn, 2002), Feature Driven
Development (Coad, Lefebvre, and De Luca, 1999), and Adaptive Software Development (Highsmith, 1997).

Perhaps Scrum is so prevalent because it has struck the right balance between overhead, and the need for working programmers to have a viable process model. Would it prevent another Therac-25 or Ariane 5-style disaster? None of these methods can prevent programmers making mistakes. The pair programming technique advocated as part of XP could help, as a programmer’s code would be scrutinised in detail as it was written, although, according to Sommerville (2011), the effectiveness of this approach is debatable.

3.6.3 Open Source software development

Open Source Software (OSS) is software for which the source code is available to the users, and which may invite contributions from the user community. Daues (2006) argues that this community-based approach is similar to the methodology of scientific research. The Open Source Initiative (www.opensource.org) defines OSS and grants licenses to compliant software.

OSS development is described by Abrahamsson et al. (2002) as a way for widely dispersed individuals to collaborate, producing software with small and frequent increments, sharing principles and practices with agile methods, although for Cockburn (2002), OSS differs from agile methodologies in philosophical, economic, and team structure models. A key benefit of the open source approach is peer review, where code inspection by many contributors uncovers more defects than in ‘closed’ development (Voightmann and Coleman, 2003). Seen in this light, OSS is another key technique in improving software quality, as reviewers could spot mistakes made by one programmer.

3.7 Video game design principles and practice

This section explores methods and guidelines for designing video games – as distinct from the software development process used to implement the design. In the context of the present research, accessibility and assistive technologies are important considerations.

Video games are software, but also much more than this – an interactive experience, which can tell a complex story and stir a deep emotional response in the player. The notions of human game playing developed by Johan Huizinga (1949) and Roger Caillois (1961) have influenced many of today's video game designers (Schell, 2008). Translating their definitions of play into the context of video games, Schell proposes that video games should have the following key qualities:

1. Entered wilfully,
2. Have goals,
3. Have conflict,
4. Have rules,
5. Can be won and lost,
6. Are interactive,
7. Have challenge,
8. Can create their own internal value, and

Blatant copies aside, every game is different, and if there is no one process suitable for developing all software (Sommerville, 2011), this goes doubly for designing games. But there is no shortage of guidelines and principles that can be applied to a game design – Schell’s 100 ‘lenses’ are the many perspectives a game designer might use when approaching his or her work.

3.7.1 Key considerations in game design
The following are some key concepts that every game design must address (Thompson, Berbank-Green, and Cusworth, 2007):

• What is the genre? Common genres are ‘shoot-em-ups’, first person shooters, platform, strategy, and puzzle games;
• Is the game single- or multi-player, or both?
• Is the game world realistic or abstract?
• Is the game 2-dimensional, 3-dimensional, or in between?
• Does the game have a strong narrative thread, with storyline, characters, and plots?
• Is the game’s structure linear or ‘sandbox’?

The term ‘sandbox’ refers to childhood play, free of constraints, and in video game terms means that the player has the freedom to explore the game world and complete game objectives in any order. Even in a sandbox game, some linearity may be imposed on the game design to ensure key story points are hit in order.

3.7.2 Play testing
Over a game’s design and development cycle, there is one technique that, if ignored, will very likely result in a poor game design: play testing. Schell (2008) gives us the indisputable rule that the more times you test and improve your design, the better your game will be.

Play testing focuses on whether the game is fun to play, and on finding issues causing frustration for players. Play tests are conducted as early as possible in the development
process, once a playable version of the game exists (Ibister and Schaffer, 2008). The ‘think aloud’ approach can be applied to play testing (Hoonhout, 2008), with the aim of finding out the following:

- Is the game an interesting and adequate challenge?
- Will it remain challenging?
- Is the game fun, and how do the different elements contribute?
- How easily can the player learn to use the game?
- For multi-player, how does social interaction develop?
- Which elements support or hinder interaction?
- Are the controls easy or do they hinder game play?

Amaya et al. (2008), conducting game usability research for Microsoft Game Studios, found that standard approaches such as Nielsen’s (1993) usability methods are not particularly well suited for games research. Amaya’s play testing approach is to recruit a ‘large’ (but unspecified) sample of gamers, who play for a specific length of time, such as one hour, then complete a questionnaire.

3.7.3 Video game accessibility guidelines

Video game accessibility is the branch of game development dealing with features that will improve inclusivity for disabled people:

‘Game Accessibility can be defined as the ability to play a game even when functioning under limiting conditions. Limiting conditions can be functional limitations, or disabilities — such as blindness, deafness, or mobility limitations.’

(International Game Developers Association, 2004)

Although researchers, hobbyists, and participants at events like Global Game Jam (www.globalgamejam.org) have spearheaded the development of games that do contain accessibility features, most commercial games ignore the issue (Game Accessibility, 2016).

In the present research, disabled participants are to play a video game, and so accessibility issues must be identified and addressed. For a game to be fully inclusive, a wide range of limiting conditions has to be catered for, with Barlet and Spohn (2012) listing the following:
• Blindness, poor vision, and colour blindness,
• Deafness and poor hearing,
• Mobility issues due to neurological disorder, and repetitive stress injury,
• Age-related issues of slower reflexes and lack of steadiness, and
• Cognitive disabilities: memory loss, attention deficit disorder and dyslexia.

Game accessibility guidelines published by the AbleGamers Foundation (Barlet and Spohn, 2012) aim to improve the number of games with accessibility features. Features are given a difficulty rating from one to three, where the easier two levels have a relatively low cost of implementation.

3.7.4 Video game assistive technologies

Practically all games require the user to be able to hold a controller, and manipulate the controls quickly and accurately in order to progress in the game. People with disabilities may be prevented from doing this due to the design of the controllers, the reaction speed required, the dexterity and accuracy required, and so on. Work has been done in making games accessible to disabled people, such as by replacing standard controllers with relatively large individual switches. These can be placed around the player allowing a game to be played using any combination of body movements. UK charity Special Effect (www.specialexpect.org.uk) has successfully used eye-tracking and bespoke controller modifications. Classes of assistive technology (AT) used for games include touch screens, eye tracking, mouse and keyboard alternatives, and switch interfaces. Eye tracking systems determine the point on the screen at which the user is focused. Most eye tracking systems work using infrared (IR) light, so as not to dazzle the user. The IR light reflected by the eye forms a bright circle at the pupil. A smaller point of IR light is reflected due to the corneal reflection, (also called the first Purkinje image). The vector from the pupil centre to the corneal reflection is processed to give the gaze direction (Poole and Ball, 2005). Sip-and-puff controllers actuate a two-position switch by inhaling or exhaling through a tube. They do not require use of upper or lower limbs but do require the user to have strong respiratory control (Arshak, Buckley, and Kaneswaran, 2006).

3.7.5 Recent developments in game controllers

This section charts the interesting developments in game controller design since 2006, the year the Nintendo Wii was released: game consoles of recent years have introduced controller hardware that is radically different to the classic joystick. Might these new controllers count as assistive technology devices, and improve usability for disabled gamers?
The Nintendo Wii could be credited with an increase in the use of video games for physiotherapy. The reason for this is the Wii remote (‘Wiimote’), which allows control of the game using arm and hand motion. The Wiimote has an infrared camera that is used to determine distance and orientation relative to the so-called sensor bar, which does not actually contain any sensors, but rather an array of infrared emitting diodes. Additionally, the Wiimote and ‘nunchuck’ accessory contain accelerometers, which can detect motion and orientation (Lee, 2008). The precision of this sensor was augmented with the release of the Wii Motion Plus in 2009 (Sung, 2011). The Wii controller can thus determine its position and orientation in 3D space, enabling new ways to control a game.

Games such as Wii Sports (Nintendo Corporation, 2006) promote upper body movement and so have been used for rehabilitation after stroke, etc. For example, Deutsch, Borbely, Filler, Huhn, and Guerrera-Bowlby (2008) report using the Wii as a complementary form of physiotherapy for an adolescent with cerebral palsy. The release of Wii Fit by Nintendo in 2008 introduced another user input peripheral, the Wii Balance Board. This is a device that the player stands on, much like scales, and it can indeed weigh the player. More interestingly, the position of the centre of gravity can be used to control a game, or in the case of Wii Fit, to provide feedback on fitness activities. Graves, Ridgers, Williams, Stratton, Atkinson, and Cable (2010) found that using the Wii Fit stimulates light to moderate intensity activity, but though preferable to sedentary behaviour, higher intensity would be required to maintain cardiorespiratory fitness.

The ‘Kinect’ controller for the Microsoft Xbox 360 console was released in 2010. The Kinect is not a hand-held controller, but a sensor array, which is placed facing the player(s), typically just below and in front of the television. The Kinect uses a range of complementary techniques to determine the position and motion of players in 3D space, allowing control of the game without holding any controller at all (Sung, 2011). The granularity of control using the Kinect’s cameras is coarse, with games requiring gross movements such as waving an arm, or moving the whole body (dancing games, for instance). The Kinect interface may therefore, like the Wii, be suitable for encouraging physiotherapy for the upper, lower, or entire body. Lange, Chang, Suma, Rizzo, and Bolas (2011) describe a balance rehabilitation game designed for the Kinect. Its advantage over the Wii interface is that the entire body of the player is monitored, eliminating the possibility of cheating the system. In addition to position and motion sensing, the Kinect offers voice recognition. This form of control has not been widely exploited as an accessibility aid but could clearly help players with limited mobility.
Sony released the PlayStation 3 in 2006, and the PlayStation Move in 2010. This consists of a camera and handheld motion-sensitive controller. The controller has a bulb at the end, which can emit light of any colour. The position and size of the bulb as seen by the camera allows tracking of the controller's position in 3D space (Tanaka, Parker, Baradoy, Sheehan, Holash, and Katz, 2012). Games using the PlayStation Move are thus as capable of providing physiotherapy as games using the Wiimote.

Overall then, we can say that game controller technology has seen some interesting recent developments, and these new controller types have been exploited for therapeutic benefit in some cases. But this is more due to the ingenuity of the rehabilitation researchers than by design: these new controllers assume non-disabled players, who can precisely aim a remote (in the case of the Wii or PlayStation Move) or precisely move their whole bodies (in the case of Kinect). Fortunately, other hardware, such as eye-tracking or sip-puff controllers can be used to enable severely disabled people to play video games. For the current research, some participants may require this kind of accessibility hardware, should they have physical impairments preventing the use of the standard joystick, mouse and keyboard input devices.

### 3.8 Conclusion

This chapter has reviewed research designs, with the aim of finding a suitable experimental method for a study. We have reviewed issues of sample size that can affect the validity of the results in a prospective study: recruitment is a difficult part of the research process, with many studies failing due to shortage of participants, and others having results with low statistical power. The single-subject research design (SSRD) solves the issue of sample size, as a time series of data is gathered for one participant. As the independent variable is changed, typically by introducing and then withdrawing some intervention, the values of the dependent variables should change correspondingly. This effect should ideally be visually apparent from a plot of the data – indeed, inferential statistical tests must be used with caution with this design, as the autocorrelated nature of the data precludes many such analyses. The downside of the SSRD is its lack of external validity.

In the present study, the SSRD is a good choice, allowing an experiment to proceed with a low number of participants. If the results show promise, the next step would be to run a larger study.

Next, we reviewed cognitive tests used in clinical practice. The aim was to find tests that could be used to detect the proposed beneficial effects of multi-player online video games, with the additional criterion that they could be implemented in software. The Mini Mental
State Examination (MMSE) and clock drawing tests are commonly used in clinical practice, but are not seen as suitable for the present study as they rely on the presence of a clinician. Stroop tests, cancellation tests, and trail-making tests have been used as broad cognitive measures in other studies, and are more suitable for the present work.

Tests of quality of life and loneliness were also considered in this chapter, and could also be used to measure the effect of playing a multi-player online video game. The Satisfaction With Life Scale (SWLS) is a well-validated test, consisting of a five-item questionnaire. The De Jong Gierveld 6-item scale is a similarly short instrument, which measures the emotional and social aspects of loneliness identified by Weiss. In the present study, it is proposed that social loneliness would be affected more than emotional loneliness.

Having identified a suitable experimental design and dependent variables, we turned to issues relevant to the implementation of a research instrument. There are many guidelines that can be used in the design of video games, but in the present study, particular emphasis is placed on accessibility issues. Human-computer interaction principles may be brought to bear, with iterative prototyping being an important technique. Assistive technologies for gamers may be used, but many accessibility improvements can be made in the game software.

Software development process models were reviewed, with an eye to finding a lightweight process that will scale up for a small team, with the intention being to make the game software open-source, enlisting the help of other programmers. Scrum is used for most agile development, while heavyweight processes have a poor track record of delivering high quality software to a schedule. Common themes of iteration and active risk management could be usefully applied to even the smallest team.

The research has taken shape. The next stage is to seek ethical approval for working with both brain-injured and neurotypical (not brain-injured) participants, to perform play testing and conduct experiments to measure the proposed beneficial effects of multi-player online video games.
4 Ethical and medical considerations

This chapter reviews ethical considerations that apply to all research, and the ethical and medical issues raised in particular by research involving video games and brain-injured participants.

All research involving human participants today is subject to ethical considerations. Research ethics comprises ‘principles and standards that, along with underlying values, guide appropriate conduct relevant to research decisions’ (Kimmel, 2007, p. 6). Modern ethical standards have been shaped by the 1947 Nuremberg Code, and the 1964 Declaration of Helsinki of the World Medical Association (Leaning, 1996). The principles of informed consent and the right to privacy are fundamental; furthermore, the Declaration of Helsinki states that it is unethical to conduct research that is badly planned or poorly executed (Greenfield, 1996).

There are three main sources of guidance to which the researcher can turn in the ethical decision making process: personal, professional, and regulatory. A researcher’s personal ethical values are shaped by his or her life experiences, and may lie on a spectrum between means-oriented (‘do no harm’) and ends-oriented (‘the ends justify the means’) (Kimmel, 2007). In addition to his or her personal value system, the researcher will be expected to adhere to standards set by their profession. Researchers at hospitals and universities are required to submit their proposed study to an ethical panel or commission, typically called an Institutional Review Board. One advantage of this is that several people judge the ethical validity of a course of action. If the group of judges all arrive at the same conclusion, the confidence in the decision is increased; if there is a conflicting decision, this indicates a possible dilemma that has been overlooked (Wueste, 2000).

Finally, legal regulations supersede personal and professional principles. Examples of such government regulations are the US Federal Policy for the Protection of Human Research Subjects (US Department of Health and Human Services, 2015), and in the EU, the 2014 Regulation 536/2014 (European Parliament and Council of the European Union, 2014).

4.1 Informed consent

Contemporary formal ethical standards can be traced back to the Nuremberg Code, arising from the Nuremberg trials of Nazi scientific atrocities. The ten-point Code introduced the concept of voluntary, informed consent, describing it as ‘absolutely essential’. Saha and Saha (2000) discuss informed consent in the context of clinical trials. Informed consent is crucial to allow the participant to decide what risks to take with his or her body. Informed
consent protects the human rights of the participant. It is too valuable a principle to be sacrificed for any anticipated research benefit.

In the context of providing medical treatment, the UK Department of Health stipulates that the following information must be provided (Department of Health, 2009):

- The benefits and risks of the proposed treatment,
- What the treatment will involve,
- The implications of not having the treatment,
- What alternatives are available, and
- What the practical effects of having, or not having the treatment will be.

In order for informed consent to be granted, individuals need to have risks and benefits explained to them. For example, in his work with locked-in patients, Paul Gnanayutham provided demonstrations to participants and their families (Gnanayutham et al., 2005).

For a person’s consent to be valid, the person must have capacity, acting voluntarily, and provided with enough information to enable them to make the decision. For a person to be capable of making a decision, they must be able to understand, retain, and weigh information. Methods of assessing this capability include exploring the person's ability to paraphrase, make comparisons, consider consequences, and apply the information given to his or her own case. Adults are presumed to be capable, unless the opposite has been shown. Where there is doubt, the capacity of the person must be assessed, with advice from specialists sought as required. A person should not be assumed unable to make any decision, even if they were unable to take a particular decision in the past.

No one can give consent on behalf of adults who are not capable of giving consent for themselves. Treatment and care may still be lawfully provided, but this must be in the person's best interests. It is generally not appropriate to perform research involving participants who do not have the capacity to give consent, if the research could instead be carried out with other people who can give or withhold consent. The only exception is where clinicians believe it would be in the person’s best interests to be involved in the research (Department of Health, 2009).

### 4.2 Privacy

The right to privacy is enshrined in major human rights codes, e.g. the United Nations’ Universal Declaration of Human Rights (United Nations, 1948), and the European Convention on Human Rights (Council of Europe, 1950). A basic principle of research ethics is that the privacy and anonymity of participants should be respected.
Issues arise when usage of a system in the field is logged. In seeking to improve usability, collecting data from real users in the field is recommended (Nielsen, 1993), to gather statistics such as how many features are used or the rate of errors, to find usability problems which are not apparent during observations, etc. Clear issues of privacy are raised: the user must be made aware if logging is being performed, and must be able to disable it.

Methodological issues may also arise due to privacy issues, as participants may be unwilling to answer questions that threaten their privacy (Kimmel, 2007, pp. 122-123). For example, in the present study, the data collected from questions about satisfaction with life and loneliness are especially sensitive. All data must be treated as confidential. Gaggioli, Gorini, and Riva (2007) caution that using virtual environments for online therapy exposes participants to risks arising from the anonymous and open nature of virtual environments. 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.

4.3 Possible negative effects of video games

Unfortunately, in addition to the therapeutic uses of video games described in chapter 2, there are health risks and other negative effects associated with playing video games. In this section we examine the medical risks and other ethical issues surrounding video games.

4.3.1 Health risks

Playing video games carries a risk of seizure due to photosensitive epilepsy (PSE) (Millett, Fish, and Thompson, 1997; Quirk, Fish, Smith, Sander, Shorvon, and Allen, 1995). Although the incidence of people with PSE is approximately 1 in 4000 (Harding and Jeavons, 1994, p. 161), the incidence of epilepsy generally (not just PSE) is much higher among people who have suffered a traumatic brain injury (TBI). About 35% of TBI patients experience a seizure, with an on-going risk of seizure in 5% of open or penetrating head injury patients. Closed head injury patients have a 1% chance of seizures (Powell, 1994, pp. 66-67).

Precautions can be taken to minimise the probability of a seizure due to PSE. The incidence of epileptic seizure correlates with the number of retinal cells stimulated, and the intensity of stimulation (Oborne, 1995, p. 158). Recommendations to reduce the likelihood of seizure are:

- Use a smaller screen, or move further away from the screen,
• For text-based systems, display light text on a dark background,
• Limit the bright areas on the screen,
• Reduce the luminance of the screen,
• Ensure the screen is placed in the central viewing area, and
• Reduce the screen-surround contrast.

These guidelines may conflict with the aim of making the display clearly visible. The ideal luminance of a display will depend on the level of ambient light in the room. For a given ambient light level, the luminance of a display must be high enough for the display to be clearly seen (Oborne, 1995, pp. 152-153).

One video game, Berzerk (Stern Electronics, 1980), has been credited with causing a real life death. In 1982, in Calumet City, Illinois, an outwardly healthy 18-year-old student played two games of Berzerk. Moments later he died of a heart attack, the autopsy revealing a congenital heart defect. The coroner speculated that the stress from playing the game was a factor, although stress from any other source would have had the same effect (Kiesling, 1982). Gwinup, Haw, and Elias (1983) measured the cardiovascular effects of playing Berzerk in 23 healthy young men. The mean heart rate and systolic blood pressure of the participants during play was significantly higher than the rate before or after.

Gwinup and team offer the explanation that the playing of video games causes the release of catecholamines – ‘fight or flight’ hormones. This was later confirmed by Koepp et al., (1998), who discovered that playing a video game caused the release of dopamine. Novice players experience greater anxiety, and a greater rise in blood pressure than for the more experienced players. In view of these results, it may be expected that video game players will experience other cardiovascular effects, such as arrhythmias; and such effects may be more pronounced in novice players.

Finally, a condition christened ‘Nintendo elbow’ was identified by Bright and Bringhurst (1992), and a variety of minor aches and pains of this type are reported by Griffiths (Griffiths, 2005). Treatment for these conditions usually consists of taking a break from playing the game in question (Griffiths and Davies, 2005).

Is it dangerous to play video games? Overall, ‘the evidence of serious adverse effects on health is rare’, although ‘frequent players are the most at risk from developing health problems’ (Griffiths and Davies, 2005).

In addition to health concerns, other ethical issues surround video games, which are examined in the following sections.
4.3.2 Video game ‘addiction’

Goodman (1990) defines an addictive behaviour as one where there is recurrent failure to control the behaviour, and continuation of the behaviour despite significant negative consequences. For many years it has been noted that someone who plays video games excessively may appear to be addicted, e.g. by Turkle (1984). Moreover, the ‘addictiveness’ of a video game is seen as a desirable quality by game players and designers, exemplified in an interview with leading game designer Sid Meier (Rouse III, 2005, pp. 26-27), Meier revealing:

‘We intend for all of our games to be fun to play and hope that they are addictive to some degree. But Civilization had a magic addictiveness that we really didn’t design, that we really didn’t anticipate.’

Griffiths and Davies (2005) have studied the question of whether or not video game addiction exists. Their opinion is that this is indeed a real condition, because six major criteria for addictive behaviour can be seen in ‘a very small minority’ of people who play video games excessively. The six criteria are:

- Salience: video games can dominate the addicted person’s thinking and behaviour;
- Mood modification: feeling a ‘buzz’ or an escape;
- Tolerance: increasing amounts of play are required to achieve the same mood altering effects;
- Withdrawal symptoms when game play is discontinued, such as tremors, moodiness, and irritability;
- Conflict: with people around the addicted player, with other activities, and internally, such as a feeling of loss of control; and
- Relapse: a tendency to revert to earlier patterns of behaviour.

Behavioural signs of addiction in adolescents that have been reported include stealing money to play arcade games or to buy game cartridges, truancy from school to play games, and not doing homework (Griffiths and Davies, 2005). Game-related crime is also reported by Loftus and Loftus (1983, pp. 109-110). A thirteen-year-old boy in Des Moines, Iowa resorted to constant burglary to fund his Pac-Man habit. In Japan, a twelve-year-old held up a bank with a shotgun, demanding only coins for arcade games. And ‘cases of children becoming prostitutes specifically to earn money for video games have cropped up in several countries’. Gentile et al. (2011) conducted a two-year longitudinal study of over 3000 schoolchildren in Singapore. They found that 9% of the children were ‘pathological’
gamers, where the excessive game use caused problems with school and sociability, and caused depression.

Arguing against the view that video game addiction exists, Wood (2007) presents case studies showing that video game ‘addiction’ could simply be categorised as bad time management. While accepting that excessive video game playing exists, he suggests that this is a symptom of an underlying problem rather than a problem in itself.

If video game addiction exists, it would seem that enabling someone to play video games would carry the risk that the player will become addicted, and this risk could conceivably be higher for people who are unable to balance game playing with a variety of other activities.

4.3.3 Violence and video games

Since the arrival of video games, concerns have been voiced over their violent nature and the possible effects on the player, echoing similar debates over violence on TV, movies, and comics. Most video games – around 80% – feature violence, with this figure rising to over 90% for games targeted at mature audiences (Smith, 2006).

Loftus and Loftus (1983) distinguish between violence to aliens and violence to other people. With regard to space games such as Defender (Williams Electronics, Inc., 1981), Galaxian (Namco, 1979), and Space Invaders (Taito Corporation, 1978), ‘the idea of defending ourselves against aliens may well be so deeply ingrained in our collective psyche that it’s futile even to worry about it’. Much more worrying to them are the ‘kill people’ variety of games, although no evidence at the time was available to demonstrate whether or not playing violent video games promoted actual violence.

Provenzo (1991, p. 65) felt that the criticism of games which emphasise violence is justified, with no distinction between games in which fairly abstract aliens are killed, and games in which humans fight each other, with the view that in any case, violent video games ‘do – at least on a short-term basis – increase the aggressive behaviour of the individuals who play them’. This appeared to be borne out when Anderson and Bushman (2001) undertook a meta-analysis of 35 studies of video game violence, and found that violent video games do increase aggressive behaviour in children and young adults. They concluded that ‘exposure to violent video games poses a public-health threat to children and youths, including college-age individuals’. This was seen as a strong view by other media researchers (Nielsen, Smith, and Tosca, 2008, p. 232), and other evidence suggests that video games are not devastating society to such an extent. While consumption of video games has risen linearly since the 1970s, the youth violent crime rate in the US remained
steady, until it began to decline in the mid 1990s, as shown in Figure 4.1 (The Economist, 2005).

![Figure 4.1. US Video game consumption vs violent crime, 1996-2004 (The Economist, 2005).](image)

Hartmann and Vorderer (2010) ask why video game violence is enjoyable in the first place, given that killing other characters could be expected to produce feelings of aversion, guilt, and disgust. Suggestions are that the player feels effective, powerful, and excited, and that the player may experience pride or euphoria due to success. The authors argue that video games provide cues that allow players to suspend their moral concerns.

4.3.4 Video games, gender, and sexuality

Consalvo (2003) has found that games from Donkey Kong (Nintendo Corporation, 1981) to Final Fantasy IX (Square Company, Ltd., 2000) have presented not only an unquestioningly heterosexual theme, but also a stereotyped view of females who invariably need rescuing by a male. Provenzo (1991) analysed the cover art of 47 popular video games, finding that they routinely portrayed women as victims, having no initiative, and dependent on men. Smith (2006) reports similar findings: of the 54% of games featuring female characters, only two featured females on the cover, both portrayed ‘provocatively’. Within the games, Smith found that female bodies in games are sometimes ‘hyper sexualised’, with unrealistically large breasts and small waists. Smith concludes that girls have fewer role models in games, and that the role models that they do have tend to be sexualized and disproportionately thin. These depictions may also affect boys’ social learning about women.
4.3.5 Video games as a corrupting influence

Some communities have banned amusement arcades on the basis of being an unwholesome environment where aggressive behaviour is encouraged. In the Philippines in 1981, President Ferdinand Marcos banned arcade games for being a corrupting influence on children (Provenzo, 1991). In 1982, the US Surgeon General, Dr C. Everett Koop, delivered a speech in which he declared video games evil, that produced ‘aberrations in childhood behaviour’, and which should not be played (DeMaria, 2007). Although this statement was not supported by any evidence, and was later retracted, possibly the sentiment was felt among some parents at the time.

Some games are unquestionably offensive by design. Provenzo (1991) notes ‘video games have a history of being sexist and racist’, citing a particularly egregious game, Custer’s Revenge (Mystique, 1982) as an example. This game attracted protests over its depictions of women and Native Americans. But deliberately offensive games are more popular and lucrative than ever, with Grand Theft Auto IV (Rockstar North, 2008) and Grand Theft Auto V (Rockstar North, 2013) – where the player assumes the role of violent criminal – breaking video game sales records (Thier, 2014; BBC, 2008). The GTA series of games has been duly criticised: Hillary Clinton in 2005 complained that ‘the disturbing material in Grand Theft Auto and other games like it is stealing the innocence of our children and it’s making the difficult job of being a parent even harder’ (DeMaria, 2007).

The video games industry has responded to concerns over unsuitable content by creating regulatory bodies. In the US, the Entertainment Software Review Board (ESRB) has been rating games since 1994. Under the ESRB scheme, there are five age-based categories and 30 content descriptors. On its release, the ESRB gave GTA IV the rating ‘M’ (mature, for ages 17+), and, perhaps fearing renewed opprobrium from Mrs Clinton, urged parents to observe their ratings. In most of the EU, the Pan European Game Information (PEGI) standard created in 2003 is used to rate games. This scheme similarly specifies age-based ratings and 6 content descriptors.

The issue of unsuitable advertising and box cover artwork has been tackled by the Advertising Review Council of the ESRB (2001), who issued guidelines for the marketing of video games. Publishers must be ‘sensitive’ in portraying violence, sex, alcohol and other drugs, offensive expression, and beliefs. Smith (2006) reports that these guidelines have been complied with on the whole, with the exception of the depiction of excessive violence.
4.3.6 **Negative effects of MMOs**

As a relatively new type of video game, massively multi-player online games (MMOs) may carry yet further risk compared to single-player games. MMOs may be more ‘addictive’ than single-player games. Griffiths, Davies, and Chappell (2004) surveyed 540 Everquest (Daybreak Game Company, 1999-2015) players, 80% of whom reported sacrificing at least one thing in their lives in order to play the game, such as sleep, work or education, socialising, or spending time on another hobby or pastime. Griffiths et al. suggest that this may be due to the social interactions, both competitive and cooperative, with other players.

Smyth (2007) conducted a randomised, prospective study into the effects of massively multi-player online role-playing games (MMORPGs) compared to other types of video game. This contrasts with other video game studies in which participants had self-selected to play MMOs. This study had 100 participants, randomly assigned to play four different video game types. Smyth found that playing MMORPGs could lead to longer game usage, worse reported health and sleep, and problems in academic and social life. Smyth does also note that most individuals do not suffer negative impact, and may derive benefit from the relationships and social interactions unique to this type of game.

4.4 **Institutional ethical approval**

To summarise, in the light of the above literature review, any experiment requiring the participants to play video games would carry the following risks:

- Seizure due to photosensitive epilepsy;
- Prolonged use of video games can cause aches and pains;
- Video games have been criticised over their use of violence and other unwholesome content; and
- Using a public MMO server would expose participants to random members of the public, which would be a cause for concern (Gaggioli, Gorini, and Riva, 2007).

Approximately 1 in 4000 people suffer from photosensitive epilepsy (PSE), although the incidence of epilepsy is much higher in TBI survivors. Minor risks are: the possibility of aches and pains, eye strain, and stress caused by extended periods of vigilance. Limiting play to short sessions minimises these risks. Other issues to consider when using video games as therapy are video game ‘addiction’, violent content, and the offensive nature of some games.
For the present research, the following measures were proposed to mitigate the risks identified.

- To minimise risk of photosensitive seizure, play sessions should be kept short.
- To minimise risk of aches and pains, and other ill effects, exposure to games will be limited to a reasonable duration.
- Using a private server eliminates the risk of contact between participants and unpleasant members of the public. Admittance to the server is by invitation only.
- The issue of any potential addiction is mitigated by short playing times.

Following the risk assessment, institutional ethical approval was granted by the Faculty of Technology Ethics Committee at the University of Portsmouth. Ethical approval was again sought and granted when the experimental protocol underwent changes. Ethical approval was confirmed by emails from Dr John Williams, Chair of the Faculty of Technology Ethics Committee at the University of Portsmouth, dated 14 March 2012 and 23 May 2014. Appendix 4 shows the Ethical Research Review Checklist for this research project.
5 Research approach

This chapter details the design of the research project. The overall approach is experimental, with the design initially based on that of a previous study (Malec, Jones, Rao, and Stubbs, 1984). That study is now described in some detail, followed by a description of the current design.


This section summarises the study ‘Video game practice effects on sustained attention in patients with craniocerebral trauma’ (Malec et al., 1984).

At the time of publication of the paper, James Malec was a psychologist at the Mayo Clinic, Rochester, Minnesota. Since then, he has been appointed Professor and Research Director, Physical Medicine and Rehabilitation, at the Indiana University School of Medicine, Emeritus Professor of Psychology at the Mayo Clinic, and Research Director at the Rehabilitation Hospital of Indiana. In 2011, Professor Malec was awarded the Robert L. Moody Prize for distinguished initiatives in brain injury research and rehabilitation, and also that year, a lifetime achievement award from the brain injury interdisciplinary special interest group of the American Congress of Rehabilitation Medicine.

The research question posed by Malec et al. was whether regular practice with a simple video game requiring sustained visual attention would enhance recovery of sustained attention in the early phases of brain injury rehabilitation. At the time, video games had been discussed as a form of cognitive rehabilitation, and there was some anecdotal evidence of success, but no controlled studies had been published. The focus on sustained attention stemmed from their observation that successful rehabilitation for ABI survivors concentrated on three key areas: attention and concentration, visual-perceptual processing, and memory.

5.1.1 Experimental design of Malec et al.

Malec et al. noted the methodological issues inherent in experimental research with brain-injured participants. Recovery is uneven and unpredictable, and spontaneous recovery can occur. The variation between individuals can be great, depending on the nature of the injury sustained. For these reasons, a large number of participants would be required for matched groups in a randomised controlled trial. The single subject research design (SSRD), described in §3.1.4, circumvents these issues, and was used in the study. Malec et al. chose to use a double crossover variant of the SSRD: half the participants started with the intervention phase, while the other half started with the normal care phase.
Participants played the video game Target Fun (Atari, Inc., 1977), shown in Figure 5.1.

The game was played on alternating weeks for four weeks. During a game week, participants played for two half-hour sessions, from Monday to Thursday. On Fridays, for all weeks, measures were administered. These were Stroop tests, letter cancellation, symbol cancellation, and a reaction time test.

In the Stroop word test, participants read columns of random colour words as quickly as possible. In the Stroop colour test, participants name random colours. In the Stroop colour-word test, the participant must name the colour of the ink in which a colour word is printed – not the word itself. The letter cancellation test required participants to cross out all 32 occurrences of the letter M, randomly distributed in a grid of 6 rows of 52 random letters. The score was the number of Ms crossed out in the three-minute time limit. In the symbol cancellation test, 36 occurrences of the Greek letter Σ (sigma) were randomly distributed in a 6 by 52 grid of random Greek letters. The score was the number of sigmas crossed out in three minutes. In the reaction time test, the time between a light stimulus and button press was measured. A ‘ready’ signal was given before the light stimulus.

Participants were patients with TBI admitted to the Rehabilitation Medicine Inpatient Service of the University of Wisconsin Hospital, selected according to the following inclusion criteria:

- Stable vital signs and post coma,
- 16 years old or older,
- Length of coma greater than 24 hours,
- Best corrected visual acuity in one eye of 20/30 or better,
• Functional use of one upper extremity, and
• Participants’ pre-study scores on the measures described had to fall within an acceptable range.

20/30 visual acuity means that one is able to see, at a distance of 20ft, what most people are able to see at 30ft; 20/20 is thus normal vision. UK standards of vision for driving stipulate a visual acuity of 20/40 or better (Great Britain, 2015). 20/30 is better than this minimum requirement for driving and ensured that participants were able to see the game.

Ten participants completed the study: eight male, two female, aged from 17 to 48. All participants had been injured within 6 months of the study. One participant did not complete the study. Two potential participants were rejected because their pre-test scores were not in the acceptable range.

5.1.2 Results and analysis of Malec et al.

Malec et al. found a mean improvement from beginning to end of the experiment for all measures, though the standard deviation was large in each case. The contribution of the intervention to the improvement was analysed by comparing scores from week to week. To do this, Malec et al. calculated a ‘percentage change towards normal’ for the test scores each week. For each measure, the calculation performed was

\[ C_n = \frac{A_n - A_{n-1}}{X - A_{n-1}} \times 100 \]

where

• \( n \) is the test session number: 1 for pre-test, 2 for end of week 1, etc.,
• \( A \) is test score,
• \( X \) is the ‘lower limit of normal’ for the measure, and
• \( C \), the result, is the ‘percentage change towards normal’.

The ‘lower limit of normal’ value, \( X \), was as follows for each measure:

• Stroop word: 88,
• Stroop colour: 65,
• Stroop colour-word: 35,
• Letter cancellation: 31,
• Symbol cancellation: 35,
• Reaction time: 35ms.

The \( C \) scores were transformed into one of three categories for two criterion levels, 10% and 25%. The three categories were:
• Positive: $C_n$ positive and greater magnitude than the criterion level;
• Negative: $C_n$ negative and greater magnitude than the criterion level;
• Zero: $C_n$ has lesser magnitude than the criterion level, regardless of sign.

These scores were calculated at the end of each week, allowing a comparison between ‘game’ and ‘no game’ weeks.

Comparisons of the distributions for the ‘game’ and ‘no game’ scores were performed using the Wilcoxon ranked sum test. No significant ($p < .05$) differences between the intervention and non-intervention weeks were found; only the reaction time test showed a significant improvement at the $p < .1$ level.

The same analysis was performed using the less stringent 10% criterion to categorise the scores for each week. Again, no significant improvement was found.

Malec et al. note that the improvement seen from the beginning to the end of the study could be expected as a consequence of natural recovery following brain injury. In the study, alternating the intervention with periods of no intervention controlled for this effect, and no significant difference was found in the measures between these phases. Malec et al. do not discount the potential for video game therapy, but recommend that video game rehabilitation strategies should be tested for their usefulness, rather than assumed to be effective.

The remainder of this chapter describes the approach developed for the present study, which takes the experimental design of Malec et al. as its starting point.

5.2 Experimental design for the present study

As noted by Malec et al., in this type of study, a randomised controlled trial (RCT) to answer the research questions would require a prohibitively large number of participants, due to the wide variation in the effects of brain injury. This issue was solved by employing a single-subject research design (SSRD) for each participant. An SSRD relies on repeated measurements of the dependent variable (Christensen, 2004), with each participant effectively his or her own control. Adding more participants does not change the nature of the design, but can provide greater external validity.

As for Malec et al., the present study uses a SSRD, split into phases in an ‘A-B-A-B’ design. The study begins with a ‘no game’ period (‘A’ phase), followed by a ‘game’ period (‘B’), which is then repeated. During ‘game’ periods, participants will be asked to play the game and interact with the other players. During the ‘no game’ phases, participants do not play the game, but do take the tests at the end of the phase. The independent variable is
whether or not the participant has played the multi-player game and interacted with other players in the preceding period. The dependent variables are the results of a battery of cognitive tests, which are described below.

Malec et al. used a double crossover design, in which half the participants started on a ‘game’ week, the others on a ‘no game’ week. If this scheme were to be used in the present study, it would seem that the number of participants playing the online game would be halved – not desirable, as social interactions within the game would be fewer. This issue could be solved by staggering the start dates of the two groups, such that the ‘game’ weeks coincided.

Each participant was asked to use the following hardware and software:

- A personal computer capable of running the game software,
- Access to the Internet,
- Any accessibility hardware as required to enable them to play games, and
- The game software, provided by the researcher via a download link.

As explained in §3.2, it was decided that participants would play on their own computer, in their own home. The rationale for this was to maximise the potential number of participants, by removing geographical location and mobility as barriers to participation. Obviously this decision does require participants to have the required equipment, which is itself a barrier.

5.3 Participant recruitment

This section describes the methods used to recruit participants to the study. It was anticipated that recruitment would be a challenge, given that Blanton et al. (2006) found ‘participant recruitment is considered the most difficult aspect of the research process’, in the context of randomised controlled trials in the arena of physical rehabilitation. They report that 60 to 80% of trials in the US are delayed by participant shortages, and that 30% of trial sites fail to recruit a single participant! To help other researchers avoid this outcome, Blanton et al. detail the steps that were taken to recruit for a trial conducted across six US sites, with a total of 222 participants.

Malec et al. recruited participants from intakes of TBI patients admitted to the University of Wisconsin Hospital, and assembled a total sample of eleven people, ten of whom completed the study. Given that a hospital admitting TBI patients constitutes a relatively rich supply of potential recruits, it was felt that in the case of the present study, ten participants could be regarded as a successful effort.
In recruiting for the present study, the researcher focused on local brain injury support
groups as sources of potential participants:

- Headway East London,
- Headway North London,
- Different Strokes East London,
- Different Strokes North London, and
- Different Strokes Central London.

Working with these local groups allowed the researcher to attend group sessions regularly:
Blanton et al. recommend personal visits with follow-up as the most effective recruitment
approach. At these support groups, the researcher:

- Gave talks about the research project,
- Handed out information sheets to the attendees, and
- Participated in therapeutic exercises and group discussions, getting to know the
group members.

A call for participants for the study was advertised on the Different Strokes UK website,
starting in June 2012, for one year. Because of the online nature of the study, as discussed
in chapter 3, participants did not need to be geographically close to the researcher, and it
was hoped that participants could be drawn from all over the UK (and even, perhaps,
internationally).

The researcher exhibited at the October 2012 Different Strokes conference, Birmingham.
At this conference, the game prototype was demonstrated on two computers, with a poster
on display. Information sheets were handed out to conference attendees. The researcher
exhibited at the UK Stroke Assembly, in Nottingham, 31 May 2013, displaying a poster,
handing out information sheets, and talking to people who had had strokes about their
experiences with video games.

In addition to recruiting brain-injured participants, neurotypical volunteers were also
recruited. The aim of this was to increase the number of players in the multi-player game:
the neurotypical participants were not required to take the tests during the study. The
researcher recruited these participants from his game programming students at a college in
East London.

All potential participants were given an information sheet providing a summary of the
experimental protocol, including potential risks, and a link to the project website which
provided full details of the project.
5.3.1 Inclusion and exclusion criteria

The inclusion criteria for all participants were as follows:

- Capable of granting/withholding consent,
- Adult,
- With no medical needs which precluded playing video games,
- Able to read,
- Able to physically control the game using an appropriate interface, and
- Having access to suitable apparatus, i.e. a personal computer with Internet access.

There was one additional criterion for brain-injured participants: to have progressed beyond the acute care and rehabilitation stage of their recovery. This differs from the Malec study, and was on the advice of Dr Jane Williams, consultant nurse in stroke care for Portsmouth Hospitals Trust (telephone conversation, 15 September 2011). The rationale is that changes in participants’ abilities are more likely to be due to the intervention, as spontaneous improvement grows less likely over time. Dr Williams’ view was that if a participant were in active rehabilitation it would be impossible to ascribe any improvements solely to video games.

The exclusion criteria for all participants follow. A participant would not be eligible for participation if any of these criteria were met:

- Unable to give or withhold consent,
- Any history of seizure due to photosensitive epilepsy,
- Any history of ill effects due to playing video games,
- Failing any other medical assessment, or on advice of carer, or
- Already regularly playing a multi-player online video game.

This last criterion, already playing a multi-player online video game, was designed to guard against the confounding effect of another, similar video game being played during the study.

5.4 Choice of measures

The experimental design calls for the participants to periodically take a battery of tests. This section describes each of the measures used, and presents the rationale for the choice. As explained in §3.2, the researcher decided that the cognitive tests should not need to be administered by a person, but rather, could be implemented in software. In addition to the
ideal characteristics of cognitive tests identified by Shulman (2000), the tests used in this study would be:

- Sensitive to a broad range of cognitive abilities, so likely to detect the proposed effects,
- Quick to administer, so as not to fatigue participants, and
- Able to be implemented in software.

Green and Bavelier (2006) point out that as all video games are different, the potential gain in perception or cognition from a given game is hard to predict. The essential ingredients of games that lead to improvements in a particular domain are not known. Therefore, broad tests of cognitive function were chosen as the measures.

It is well known that playing action video games improves the player’s reaction time (Green and Bavelier, 2006; Griffiths, Kuss, and Ortiz de Gortari, 2013). The current game is not an action game, but if there were time-sensitive aspects to it, and the game was played in real time, it could be expected that an improvement in reaction time would be seen. If no improvement in reaction time was seen, this could indicate that a parameter of the experiment needed adjustment – e.g. the session duration had not been long enough to cause any improvement. A reaction time test was included as a measure for this reason, serving as a control.

Following Malec et al., Stroop tests were included as measures. The Stroop colour-word test is a commonly used measure of executive function, as reviewed in chapter 3. It measures the ability to shift cognitive set, and is believed to measure cognitive inhibition (Homack and Riccio, 2004), the ability to suppress a learned response in favour of an unusual response. This broad-based cognitive measure was seen as appropriate for the study. Some comparison with the Malec study would then be possible, although the present study design is not intended to be an exact replication.

The letter and symbol cancellation tests used by Malec et al. were not used in the present study. They were implemented in software and piloted, but ultimately rejected because of the relatively long time it took to complete them: up to 6 minutes. This is a problem given that the total time a participant can spend using the game and tests is limited. This time limitation is due to fatigue, and the need to minimise any possible health risks from prolonged exposure, as detailed in chapter 4. Additionally, cancellation tests primarily measure spatial neglect, and are not a broad cognitive measure (Rorden and Karnath, 2010), so were considered expendable, and were replaced with trail-making tests. Trail-making tasks are widely used to measure cognitive flexibility and attention, as well as visual capabilities (Zakzanis, Mraz, and Graham, 2005).
5.5 Summary

This chapter has described the approach taken for this research. A prior study (Malec et al., 1984) into the potential benefits of a simple single-player game was used at the starting point to develop the design for the present study. The overall approach is experimental, using a single-subject research design. This choice of design is appropriate given an expected low number of participants, but external validity of the results is correspondingly limited. A range of recruitment methods was used, based on the recommendations of Blanton et al. (2006). The researcher focused most on personal visits to brain injury support groups in North and East London.

The main differences between the current approach and that of Malec et al. are:

- The research setting: participants use their own computer in their own home;
- The length of time since participants’ brain injury; and
- The measures: letter and symbol cancellation tests were rejected in favour of trail-making tests.

The measures chosen for the study are broad-based cognitive tests:

- A reaction time test,
- Stroop colour, word and colour-word tests, and
- Trail-making tests ‘A’ and ‘B’.

The next chapter describes the software implementation of these tests, and the multi-player online game itself.
6  Design and development of prototype multi-player online game with integrated tests

This chapter describes the design and development of the research instrument. The cognitive tests identified in chapter 5 were implemented in software, integrated into a prototype multi-player online video game.

6.1  Software development methodology

Chapter 3 surveyed software development processes, principles of video game design, and principles of HCI design. In the current research, the methodology must cover these additional considerations:

• Identification and architecture of the software components that make up the system, and the communication between them;
• Consideration of build or buy options, and relatedly, deciding whether the software should be open or closed source;
• Choosing technologies, tools, languages, and libraries;
• Deciding which client platforms to target;
• Choosing a vendor for hosting services; and
• Implementing project infrastructure, such as source code control and bug tracking.

6.1.1  Software components

There are two main software components: the game client with integrated tests, with which the participants interact; and server-side software for mediating the online game, user administration, storing test data, and reporting the results. The software system as a whole comprises the following components:

• Client game executable, downloaded from server;
• Client game data (the game ‘assets’);
• Client installer (installs client executable, base resources, DLLs);
• Server-side scripts, to implement game mechanics and to record test scores;
• Server-side database;
• Administration web pages, for adding players and groups;
• Informational web site for potential players and participants;
• Notification system, which notified the researcher when participants logged into the game;
• Server-side scripts to generate test score reports; and
- Software to generate plots and analyses from the test score reports.

6.1.2 Bespoke or off-the-shelf software

For the components that are unique to the experiment, and reflect the experimental design, the only option is to develop bespoke software. But for other software components, there may be the option of using pre-existing software. This option is most plainly evident in the choice of game software: the game played by the participants could be an existing, ‘off-the-shelf’ game, or could be a bespoke game developed expressly for this research.

Developing a bespoke game allowed for all aspects of the game experience to be customised to the requirements of the experimental design and the special needs of the participants. 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.
- It allows us to have single- and multi-player versions of the game, if the experiment calls for a comparison between these two modes. If two completely different games were used for these two experimental phases, there would be much more scope for confounding issues.
- Individual features could be added and changed, potentially allowing their effects to be measured, addressing Research Question 2.
- Accessibility features can 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 a player's in-game position could be logged, so that the movements of a player can be replayed and analysed.

6.1.3 Software development process model

Bates (2001, p. 230) states that ‘the best lifecycle model for action games probably falls somewhere between the modified waterfall and the iterative prototype’. This project used the staged release model proposed by McConnell (1998), which fits this description well. Major architectural decisions are made at the start of the project, before a cycle of iterations progressively adds more features to the software.

All the software developed is open source, i.e. freely available to users. One reason for this was to enable other researchers to use any parts of the software in similar studies. Less altruistically, the researcher also hoped to enrol other volunteer programmers to help with
the software development, but this did not happen in practice. The source code, build scripts, assets, etc. were initially all stored in a Subversion repository hosted by Google (http://code.google.com). In 2015, Google announced the discontinuation of their hosting service, and the repository was migrated to a Git repository hosted on GitHub (http://www.github.com).

A bug-tracking database was set up, and testers were encouraged to file reports of bugs they found. The bug database used was that provided as part of Googlecode project hosting. In practice it was found that participants did not file any bug reports.

6.1.4 Deciding which client platforms to support

The client software was initially envisaged as being multi-platform, enabling as many participants as possible to take part by minimising the need to own a particular type of hardware. To facilitate this, technologies for developing client software were chosen with portability in mind. Another factor was the aim of speeding up development by reusing existing code. The ambition of releasing on multiple platforms was eventually scaled back, in order to reduce the time required to develop and support the game. It was found that all the participants had a PC running Microsoft Windows, so this was the only platform on which client software was released. Nevertheless, the codebase remained essentially cross-platform, as it was written in a portable language (C++) and used libraries which are available on multiple platforms. This may be helpful in the future, where it may be possible to support more devices.

6.2 Administration system

The administration system developed lets different participant groups take tests and play the game, according to per-group schedules. This design allows for flexibility in the experimental design, such as two groups in a double-crossover design, (one ‘A-B-A-B’, the other ‘B-A-B-A’), or multiple groups starting at different times, i.e. a multiple baseline SSRD. The basis of this design is a ‘research group’, which has:

- A list of participants,
- A repeating pattern of dates on which the game can be played, and
- A repeating pattern of dates on which the tests should be taken.

The administration tasks that can be performed using the administration system are:

- Create a new ‘research group’, setting when participants should play, and when to take the tests.
- Add a new player (participant), assigning the player to a research group.
The administration system was implemented as an html page, which sent requests to Perl scripts to write to the database. Figure 6.1 shows the administration web page.

![Administration web page](image1)

Figure 6.1. Administration web page.

The client requested the experimental schedule from the server on start up, and this was displayed to the player in calendar form, shown in Figure 6.2.

![Calendar](image2)

Figure 6.2. The game client displayed a calendar to the participant, showing when to play the game and when to take the integrated tests.
6.3 Game design and development

The game developed is a multi-player online adventure game, intended as a much-simplified example of a multi-player online game (MMO). This genre was chosen because of its popularity – there are millions of MMO players. Many MMOs are available ‘off the shelf’ (or more accurately now, available to download). The researcher considered using such a ready-made game as the instrument. It was decided that it would be advantageous to create a simplified game whose key mechanic was to interact with other players altruistically. The intention was that such a simplified game would allow the research questions to be investigated with fewer confounding factors than would otherwise be the case. The intention was not to create a game specifically for use by brain-injured players, although the simplicity of the game design was intended to provide a widely accessible game experience.

6.3.1 Game design framework

This section presents the overarching ‘philosophy’ guiding the design and development of the prototype game. This provided a framework for resolving the many decisions that had to be made as part of software construction.

The game was designed to foster altruistic and collaborative behaviour by placing the players in a shared, adverse situation. Andras, Lazarus, and Roberts (2007) note that in many contexts, cooperation between individuals improves as adversity and uncertainty increases in the environment. Ray (2004) found that games that place players in a ‘socially significant’ situation heighten the emotional involvement for players, increasing engagement. The aim of the game design was to exploit these concepts, and so have the effect of improving co-operation, counteracting feelings of loneliness and isolation.

The game design and development was guided by the following principles, driven by the above design concept and the ethical issues reviewed in chapter 4.

1. The game should encourage collaborative play, because this aspect of MMOs is reported to be players’ favourite feature, and, the researcher felt, the most promising route to counteracting social loneliness.

2. The game should encourage communication and social interactions between the players, as this provides the proposed beneficial effect.

3. The game must be non-violent – or at least, there should not be violence between the players, as this would contradict point 1.
4. The game should be enjoyable to play or even simply to ‘visit’, with the intention that this would strengthen the feeling of belonging to a community.

5. Referring to the potentially negative effects of video games surveyed in chapter 4, the game should of course not exhibit the negative traits discussed, and should not cause nor exacerbate health issues.

6. The game and tests must be accessible for all the participants. This means the game design should be as simple as possible, with a straightforward mechanic, simple controls, and having a clear visual style, following accessibility guidelines (Barlet and Spohn, 2012) where applicable.

### 6.3.2 Game design process and evolution

The approach taken to game design was to create a basic game prototype, then present it to play test participants for feedback, make changes, and iterate. Gnanayutham and George (2009) describe a process for developing neurorehabilitation software. Following their recommendations, the first few iterations of the game design process took place with neurotypical volunteers. The first group of participants in this process were the researcher’s family, then neurotypical colleagues and friends at the researcher’s place of work. Play testing with brain-injured participants followed, with the first play test session taking place with six people from Headway East London on 17 August 2012. None of these play test participants were participants in the later studies of 2013 and 2014.

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. The visual style shown in Figure 6.3 (overleaf) was rejected by the participants, as ‘too cute’, and was replaced with the style shown in Figure 6.4 (overleaf).

The next opportunity for feedback from potential participants was at the Different Strokes conference in Birmingham on 13 October 2012. The researcher exhibited the game prototype, with the objectives of recruiting participants, and obtaining feedback on the game design. From the conversations of that day, it became apparent that a 3D game would exclude some brain-injured people who might feel nausea or be unable to process the visual display. The visual style of the game was changed to 2D following this feedback.

In the final design, players inhabit a shared environment containing food, treasure, and harmful enemies, which deplete the players’ health. Eating food restores health, but players can only eat food given to them by other players, and are thus reliant on each other to stay alive. Players can communicate with each other by typing text into a shared chat window. These messages are broadcast to all the players. Text is also displayed to indicate the location of other players and their recent actions. This design reflects the game design
philosophy and goals of the project. Figure 6.5 (overleaf) shows the final 2D game. Table 6.1 (overleaf) lists the major game design features that were implemented over the course of the game’s development, with their intended purpose. Some of these features were later discarded or superseded.

Figure 6.3. Initial visual style.

Figure 6.4. Second visual style.
Figure 6.5. Final game design.

Table 6.1. Game design features, both included and discarded, and their intended purpose.

<table>
<thead>
<tr>
<th>Feature</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single player mode</td>
<td>Allow for experimental designs which compare single- and multi-player versions of the same game. Single-player mode was not used in the final design, but each game mechanic has a possible single-player equivalent.</td>
</tr>
<tr>
<td>Player timing ability or fast reflexes must not be required to play</td>
<td>A game mechanic based on reaction time or motor skill would be likely to exclude some participants.</td>
</tr>
<tr>
<td>Collaboratively move ‘building blocks’ to solve puzzles or problems; use the blocks to create in-game buildings and other features</td>
<td>This was the original game mechanic envisaged, intended to allow players to solve puzzles and create game content collaboratively. This was later dropped when the game changed from 3D to 2D.</td>
</tr>
<tr>
<td>3D graphics</td>
<td>The initial design was 3D, as this allows for the ‘building block’ game mechanic.</td>
</tr>
<tr>
<td>Player-to-player chat</td>
<td>The chat feature was initially private between pairs of players. This was changed to a ‘broadcast’ chat so all players would see conversations, as it was decided that this would improve the feeling of being part of the community.</td>
</tr>
<tr>
<td>Players each have a ‘home’ area/piece of land</td>
<td>As part of the initial ‘building blocks’ game idea, this would be the player’s space in which to build on their own. Other spaces would be shared to allow collaboration.</td>
</tr>
<tr>
<td>Feature</td>
<td>Purpose</td>
</tr>
<tr>
<td>----------------------------------------------</td>
<td>----------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Turn-based fighting</td>
<td>A turn-based battle system was considered and rejected as the researcher felt a more cooperative game mechanic would better address the research questions.</td>
</tr>
<tr>
<td>Animal avatars</td>
<td>The game initially used animal instead of humanoid player characters. This was intended to remove any similarity between a player’s physical appearance and their avatar. Users rejected this idea.</td>
</tr>
<tr>
<td>Avatar selection/customisation</td>
<td>Players were initially offered a choice of different avatars. User feedback revealed the importance of customisable avatars, to make them look like the player, or other well-known characters.</td>
</tr>
<tr>
<td>Game mechanic: collect fuel cells, bring to rocket</td>
<td>This game idea was piloted as a way to encourage players to work together. This was rejected, as, although the players collaborated, the mechanic did not incentivise the players to interact with each other.</td>
</tr>
<tr>
<td>Simple point-and-click to move control</td>
<td>The game can be played entirely by moving the mouse and clicking the left mouse button. The intention was to make the game simple to ‘pick up and play’. This control system was used in the final game design.</td>
</tr>
<tr>
<td>Reliant on other players to eat to regain health</td>
<td>This mechanic provided the incentive for players to interact with and help each other.</td>
</tr>
<tr>
<td>Accessibility features</td>
<td>Soft keyboard, text-to-speech, and joystick control were added to make the game more accessible.</td>
</tr>
</tbody>
</table>

It is important for players to be able to choose an avatar that reflects them physically (Ray, 2004). In conversations with play test participants it was found that some participants wanted their avatars to look like them, or to look like some other well-known character. So customisable avatars were seen as an important software requirement. Figure 6.6 (overleaf) shows the avatar customisation screen.

During the development process, feedback from potential participants guided the inclusion of the following accessibility features.
• Text-to-speech was incorporated for all in-game instructions;
• A soft keyboard allowed for mouse-only operation, shown in Figure 6.7 in an early 3D prototype;
• A joystick could be used to control the game, instead of a mouse.

Figure 6.6. Avatar customisation screen, showing two different characters.

Figure 6.7. Soft keyboard.
6.4 Implementation of the cognitive tests

This section details the implementation of the cognitive tests used in the study. The tests are administered by software, with the intention that the physical presence of the researcher would not be required. The cognitive tests are not only computerised, but are also an integrated 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.

All the cognitive tests had a practice mode. In this mode, an arrow pointed to the correct button or target to click.

6.4.1 Implementation of the reaction time test

Participants are asked to click an on-screen button as quickly as they can after a prompt is given. Figure 6.8 (overleaf) shows the different states of the reaction time test. After presenting instructions, the client software waits for a random time period of a few seconds. The stimulus is then presented: ‘GO!’ text is shown, and a buzzer sound is heard. After this stimulus, the participant should click the button as soon as possible. Clicks before the stimulus cause the time delay to continue for longer.
6.4.2 Implementation of the Stroop tests

Figure 6.9 (overleaf) shows the implementation of the Stroop colour test. In the PEBL (Mueller and Piper, 2014) Stroop test 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. The button order does not change, following the PEBL Stroop test design, where the key mapping does not change. There is a time limit of 45 seconds, following Malec et al. (1984). Figure 6.10 (overleaf) shows the Stroop word test implementation, and Figure 6.11 (overleaf) the Stroop colour-word test. The rectangle or word moves onto the screen from the left hand side each time. This is so that if the same random colour appears in succession, there is a visible indication that this is a new item, requiring a new button click.
Figure 6.9. Stroop colour test. The 'green' button is the correct choice in this case.

Figure 6.10. Stroop word test. The 'red' button is the correct choice in this case.

Figure 6.11. Stroop colour-word test. The 'blue' button is the correct choice in this case.
6.4.3 Implementation of the trail-making tests

There are two varieties of the trail-making test. The first is to select the circles in ascending numeric order. The second test is to select the circles in an alternating pattern of ascending numbers and letters. In the alternating version of the trail-making test, the participant starts on ‘1’, then moves to ‘A’, then ‘2’, then ‘B’, etc. These tests are shown in Figure 6.12 (test ‘A’) and Figure 6.13 (test ‘B’).

In the current implementation of the trail-making tests, the participant uses the mouse, moving the cursor over the target circles. Sequential target circles are placed closely to each other, as far as possible. This follows the design of the ‘Connections’ variant of the trail-making test, (Salthouse et al., 2011), which attempts to minimise the influence of visual search and hand movement, the better to focus on the other cognitive processes believed to be involved in the test.

Figure 6.12. Sequential number trail-making test (test ‘A’).

Figure 6.13. Alternating number-letter trail-making test (test ‘B’).
6.4.4 Letter cancellation tests

Figure 6.14 shows the letter and symbol cancellation tests implemented, which followed the description of the cancellation tests in Malec et al. (1984). These tests have a three-minute time limit, so taking both tests takes six minutes. As discussed in §5.4, these tests were piloted but it was ultimately decided to remove them from the software. This was because of the relatively long time the tests took, and their specific focus as a tool for diagnosing spatial neglect.

![Figure 6.14. Letter and symbol cancellation tests.](image)

6.4.5 Recording and reporting test scores

The test scores for a participant were stored on that participant’s computer as text files, and periodically uploaded to the server until an acknowledgement was received. This design ensured that test scores were not lost due to a crash or other unexpected software behaviour. Some of the results from the tests could potentially be unreliable. For example, in a Stroop test, a participant may simply click the same button every time (out of interest, it turned out that this was never the case). To detect this kind of ‘bad’ data, all button presses were recorded, rather than just a count of correct and incorrect choices.
In the game software, participants could view their test scores as a chart, shown in Figure 6.15.

![Figure 6.15. Client results screen.](image)

### 6.5 Summary

This chapter has detailed the design of the research instrument, a prototype multi-player online video game, with integrated tests. The overall design allows a participant to play the game and take the tests without requiring the presence of the researcher. The game was designed to encourage cooperation and altruistic behaviour, putting the players in a shared, adverse situation where food had to be shared. The following cognitive tests were implemented:

- Stroop colour, word, and colour-word tests,
- Trail-making tests ‘A’ and ‘B’,
- A reaction time test, and
- Letter and symbol cancellation tests, although these were later removed from the study.

An administration system allowed for the creation of new user accounts and groups of participants, with associated dates for tests and play sessions. This schedule was displayed in the game client for the participant to view, as were the test results for that participant, which were uploaded from the game client to a server for the researcher to collect.

The development process was iterative, with firstly neurotypical, then brain-injured participants contributing feedback and participating in play testing. Due to this input, the game design was refined considerably over the development cycle.

The next chapter describes the experiment of 2013 in which this software was first used.
7  Experiment #1: Measuring cognitive improvement

This chapter details the experiment carried out over August-September 2013. The hypothesis under test was

**H1**: Multiplayer online video games contribute to an improvement in cognitive function.

The experimental design described in chapter 5 was followed, using the software implemented as detailed in chapter 6.

7.1  The participants

Brain-injured participants were recruited from two organisations: Different Strokes (www.differentstrokes.co.uk), and Headway (www.headway.org.uk), both UK organisations that support survivors of stroke and other brain injuries.

7.1.1  Recruitment activities

The strategy described in chapter 5 was followed, with the aim of recruiting a small number of brain-injured people: ten would have been regarded as a successful recruitment effort. As described previously, the researcher followed recommendations from a previous study, which successfully recruited 222 participants (Blanton et al., 2006). A list of activities undertaken by the researcher to recruit participants is shown in Table 7.1 (overleaf). In addition, the researcher advertised for participants on the Different Strokes website, which turned out to be more successful than the personal approach recommended by Blanton et al. The advertisement took the form of an item in the list of research projects maintained on the Different Strokes website, and periodic mentions in the Different Strokes Twitter stream. These linked to the participant information and consent form for the project.

7.1.2  Participant details

Three brain-injured participants completed the consent form and participated in the study. One participant was recruited as a result of the visits made by the researcher to Different Strokes Central London group. The other two responded to the advert on the Different Strokes website and Twitter feed. These three participants were allocated participant IDs 13, 15, and 22. Participant IDs were allocated sequentially over the course of the entire research project, and are not consecutive for the participants in this experiment. Basic information for these three participants is given in Table 7.2 (p. 94).
### Table 7.1. Recruitment activities for 2013 study.

<table>
<thead>
<tr>
<th>Date</th>
<th>Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>13 Oct 2012</td>
<td>Exhibited at Different Strokes Conference, Birmingham. (Poster and game demo running on two computers.)</td>
</tr>
<tr>
<td>1 February 2013</td>
<td>Gave talk at Different Strokes East London group.</td>
</tr>
<tr>
<td>8 Feb 2013</td>
<td>Different Strokes East London again, got some consent forms.</td>
</tr>
<tr>
<td>13 Feb 2013</td>
<td>Visited Different Strokes North London. Gave out info forms.</td>
</tr>
<tr>
<td>19 Feb 2013</td>
<td>Visited Different Strokes East London – gave out more info forms.</td>
</tr>
<tr>
<td>6 March 2013</td>
<td>Visited Different Strokes North London group again. Gave out more info forms.</td>
</tr>
<tr>
<td>15 March 2013</td>
<td>Visited Different Strokes East London.</td>
</tr>
<tr>
<td>17 May 2013</td>
<td>Visited Different Strokes East London.</td>
</tr>
<tr>
<td>31 May 2013</td>
<td>Exhibited poster at Stroke Assembly, Nottingham.</td>
</tr>
<tr>
<td>21 June 2013</td>
<td>Visited Different Strokes East London, showed game prototype to a group of ~10 group members.</td>
</tr>
<tr>
<td>13 July 2013</td>
<td>Went to Headway East London open day, then visited Different Strokes Central London group.</td>
</tr>
<tr>
<td>17 July 2013</td>
<td>Attended Headway North London meeting.</td>
</tr>
<tr>
<td>26 July 2013</td>
<td>Visited Different Strokes East London.</td>
</tr>
<tr>
<td>27 July 2013</td>
<td>Visited Different Strokes Central London. Got consent forms from 2 participants, and gave out more info forms.</td>
</tr>
</tbody>
</table>
More detailed information about the brain-injured participants now follows. Participant 13 was (and indeed, is) a chartered surveyor, owning his own company. In 2010 he suffered a series of brain injuries, ultimately diagnosed as cerebral vasculitis and multiple strokes. Describing this period, he writes:

‘The simple description of cerebral vasculitis is the complete breakdown of my immune system although the doctors had a lot of problem identifying it early doors and I had 5 nearly 6 months in hospital being tested for every brain injury that they could think of as I gather that it is an extremely rare condition and they almost have to discount every other brain injury first. The illness caused me to have over 30 strokes, the majority of which were when I was in hospital but they believe that I was having earlier ones beforehand that may have been TIA’s [Transient ischaemic attack] rather than full strokes. They certainly didn’t give me much chance of even leaving the hospital alive in those earlier days.’

Since his injury, participant 13 has not been able to work, as both his speed of thought and writing ability have been affected. In his own words,

‘Unfortunately I haven't worked in the past 5 years having tried to return to my old line of work at first but very quickly realised that my brain doesn't work the way it used to. My speed of thought is very poor, my writing is pretty terrible and I get confused easily if I don’t have plenty of time to think about things.’

Participant 13 has been affected physically by his brain injury, with impairment to his sense of balance:

‘My horrendous balance is possibly my only real physical issue. I have fallen and even hospitalised myself on several occasions but in general once I get walking and get my balance sorted I can actually then run fairly long distances and have even ran a half marathon post hospital but primarily because it is the one sport that now I

Table 7.2. Brain-injured participants.

<table>
<thead>
<tr>
<th>Participant ID</th>
<th>Brain injury details</th>
<th>Sex</th>
<th>Age</th>
<th>Prior experience of multi-player online video games</th>
</tr>
</thead>
<tbody>
<tr>
<td>13</td>
<td>Strokes, cerebral vasculitis, 2010-2011</td>
<td>M</td>
<td>45</td>
<td>None</td>
</tr>
<tr>
<td>15</td>
<td>Stroke, 2010</td>
<td>M</td>
<td>55</td>
<td>None</td>
</tr>
<tr>
<td>22</td>
<td>Stroke (haemorrhage), 2001</td>
<td>F</td>
<td>56</td>
<td>None</td>
</tr>
</tbody>
</table>
can still do reasonably. I also play walking football now which I love, but having been a semi-professional footballer in my prime my new ability is absolutely appalling but I still have fun!”

Finally, participant 13 reports that he had excellent directional awareness pre-injury, but now needs to think hard about directions, particularly in a new area. Participant 15 attended a secondary modern boys’ school, leaving in 1974 with no qualifications. In 2010 participant 15 suffered a stroke, recalling: ‘I had a brain bleed and a blockage, needed resuscitating twice while in hospital.’ He has been able to return to work as a security team leader at a large power station: ‘I was off work for 18 months and went back a day at a time, it took me about 6 months before I came back full time.’ Physically, participant 15 has been left with stiffness in his right leg, and weakness in his right arm, but reports that his problems are mainly cognitive:

‘Cognitively, well, understanding things, words meanings, information. I have trouble concentrating on things, easily distracted; of course the memory is useless. Difficulty processing information. I cannot read a book or newspaper but have no trouble reading a computer screen.’

Asked if he has problems now finding his way around, participant 15 responded,

‘Only slightly now I try and avoid places where I haven’t been before unless I am with somebody, because I can get disorientated and distressed.’

Participant 22 worked as a nurse at a London hospital, and was awarded an MBE for services to health care in 1998. She suffered a stroke in 2001, severely affecting her mobility, following which she used a wheelchair. She was unable to return to work but became coordinator of the Different Strokes Central London group.

In addition to the brain-injured participants, thirteen non-brain-injured participants were recruited, who were students on a games programming degree course, taught by the researcher. These participants had mostly had experience of multi-player online video games, in contrast to the brain-injured participants. Twelve of these thirteen were male, with ages ranging from 20-32. The role of these participants was to populate the multi-player game, not to take the tests.

Prior to taking part, all participants were given an information sheet, shown in Appendix 1, and were able to ask the researcher questions. A website provided additional information. All participants were physically able to sign the consent form.

7.2 Method

This section describes the experimental method designed to test the hypothesis. The participants were to engage in playing the prototype multi-player online video game,
designed to encourage cooperation and collaboration, and periodically take the tests built into the game software.

On five consecutive Fridays, the participants took a short battery of cognitive tests. The participants were asked to play the multi-player online game on each day from Monday to Thursday, after the second and fourth Friday of tests. Participants used their own computer, in their own home (or possibly elsewhere, in the case of laptops). The intention was to maximise the number of participants by removing geographic location and mobility as barriers.

The measures recorded in each test session were as follows:

- A reaction time test,
- Stroop colour, word and colour-word tests, and
- Trail-making tests ‘A’ and ‘B’.

The rationale for the inclusion of each of these measures is given in chapter 5, and implementation details are given in chapter 6.

Participants were not given explicit training on how to play the game. It was found during play test sessions that participants were comfortable with the point and click mechanic. The cognitive tests, which are also part of the game software, had a practice mode for each test. This allowed the participant to test his or her understanding of the test.

Malec et al. (1984) used a double crossover design, where the participants were split into two groups, one starting on a ‘game’ week, the other on a ‘no game’ week. The experiment and software of the present study were designed to allow this scheme. The two groups would start one week apart, the ‘no game’ group first, so that the two groups’ game-playing weeks coincided. The administration software was designed around the idea of multiple participant groups, to allow for this staggered start. This would also extend well to multiple-baseline variants of the SSRD. In the event, the double crossover scheme was not used in the present study. The reason was that the low number of participants would have made it impossible to ascribe any differences between the groups to this variable.
7.2.1 Schedule

Participants were asked to engage in the experiment from 23 August 2013 to 20 September 2013, as detailed in Table 7.3. On five consecutive Fridays, the participants were asked to take a series of cognitive tests, and to play the multi-player online game on each day from Monday to Thursday, after the second and fourth Friday of tests.

Table 7.3. Schedule for 2013 experiment.

<table>
<thead>
<tr>
<th>Friday</th>
<th>Saturday</th>
<th>Sunday</th>
<th>Monday</th>
<th>Tuesday</th>
<th>Wednesday</th>
<th>Thursday</th>
</tr>
</thead>
<tbody>
<tr>
<td>23 August</td>
<td>24</td>
<td>25</td>
<td>26</td>
<td>27</td>
<td>28</td>
<td>29</td>
</tr>
<tr>
<td>2013 Take</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>tests</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>31</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Take tests</td>
<td></td>
<td>September 2013</td>
<td>Play game</td>
<td>Play game</td>
<td>Play game</td>
<td>Play game</td>
</tr>
<tr>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
</tr>
<tr>
<td>Take tests</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>14</td>
<td>15</td>
<td>16</td>
<td>17</td>
<td>18</td>
<td>19</td>
</tr>
<tr>
<td>Take tests</td>
<td></td>
<td></td>
<td>Play game</td>
<td>Play game</td>
<td>Play game</td>
<td>Play game</td>
</tr>
<tr>
<td>20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Take tests</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Four phases were scheduled. As Christensen (2004) explains, any SSRD must consist of at least three phases: an ‘A-B-A’ pattern in which a baseline is established in the first phase, the intervention is introduced in the second phase, and then withdrawn in the third. This is so that any change seen in the second phase will revert to the baseline, and can therefore be attributed to the intervention, rather than unknown causes. In the present experiment, four phases were planned in an ‘A-B-A-B’ pattern, to give a second opportunity for any effects of the intervention to develop.

Each of the four phases was one week in duration. It was not known if one week would be a sufficient time for the proposed beneficial effects to appear, but this duration was used in the prior study used as a reference (Malec et al., 1984). The rationale, therefore, was to use the same duration as that earlier study, and compare results.
Lerdal et al. (2009) state that fatigue is a common complaint following a stroke, but that knowledge regarding post stroke fatigue is limited. In the current study, session length was intended to be kept as short as possible, to minimise fatigue. This had to 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. In an unstructured interview with one potential participant, it was found that 15 minutes might be the maximum acceptable duration. It was recognised that a small number of participants would probably mean that few players would be online at the same time. This problem was addressed by asking all participants to log in at the same time when playing. The consent form included a section where the participant indicated convenient times.
7.3 Results

For each brain-injured participant, for each measure, a time series of test scores were recorded. The tests were taken weekly, for a total of five times, shown in Table 7.4.

Table 7.4. Test session dates.

<table>
<thead>
<tr>
<th>Week</th>
<th>Date</th>
<th>Note</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>23 August 2013</td>
<td>Start of the study</td>
</tr>
<tr>
<td>2</td>
<td>30 August 2013</td>
<td>After the first ‘no game’ week</td>
</tr>
<tr>
<td>3</td>
<td>6 September 2013</td>
<td>After the first ‘game’ week</td>
</tr>
<tr>
<td>4</td>
<td>13 September 2013</td>
<td>After the second ‘no game’ week</td>
</tr>
<tr>
<td>5</td>
<td>20 September 2013</td>
<td>After the second ‘game’ week</td>
</tr>
</tbody>
</table>

The expected results are as follows:

- No change from week 1 (baseline) to week 2 (‘no game’ week),
- Improvement from week 2 to week 3,
- No improvement from week 3 to week 4, and
- Improvement from week 4 to week 5.

Participant 22 was taken ill, and did not complete the final round of tests. Other than that, there are a total of 15 results collected for each week, for each participant. They are:

- The number of correct and incorrect choices for each of the three Stroop tests;
- Three results from the reaction time test, as it is given three times; and
- The number of correct and incorrect choices, and the time taken, for each of the two trail-making tests.

As noted above, participant 22 did not participate in the final week of the experiment due to ill health. Her results can still be analysed as they span the first, baseline phase, the intervention, and the return to normal care conditions, i.e. the ‘A-B-A’ phases of the SSRD.
7.3.1 Stroop tests

This section presents the Stroop test results for all the participants. The results for the Stroop tests are the number of correct and incorrect choices made in 45 seconds. This time limit does not vary. Each test yields a number of correct and incorrect choices. The percentage of correct choices is calculated as

\[
\text{Percentage correct} = \frac{\text{correct}}{\text{correct+incorrect}} \times 100.
\]

Table 7.5 shows the results for the Stroop word test, for the three participants, for each of the five test dates listed in Table 7.4 above. Table 7.6 shows the Stroop colour results, and Table 7.7 shows the Stroop colour-word results. For each of these tables, the results for weeks 3 and 5 are expected to show improvement, as they follow a ‘game’ week. Figure 7.1 - Figure 7.3 (overleaf) show these results graphically, for each participant individually.

Table 7.5. Stroop word results.

<table>
<thead>
<tr>
<th>Participant ID</th>
<th>Stroop word % correct week 1</th>
<th>Stroop word % correct week 2</th>
<th>Stroop word % correct week 3</th>
<th>Stroop word % correct week 4</th>
<th>Stroop word % correct week 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>13</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>15</td>
<td>76</td>
<td>83</td>
<td>83</td>
<td>81</td>
<td>73</td>
</tr>
<tr>
<td>22</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>

Table 7.6. Stroop colour results.

<table>
<thead>
<tr>
<th>Participant ID</th>
<th>Stroop colour % correct week 1</th>
<th>Stroop colour % correct week 2</th>
<th>Stroop colour % correct week 3</th>
<th>Stroop colour % correct week 4</th>
<th>Stroop colour % correct week 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>13</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>15</td>
<td>67</td>
<td>73</td>
<td>71</td>
<td>75</td>
<td>68</td>
</tr>
<tr>
<td>22</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>

Table 7.7. Stroop colour-word results.

<table>
<thead>
<tr>
<th>Participant ID</th>
<th>Stroop colour-word % correct week 1</th>
<th>Stroop colour-word % correct week 2</th>
<th>Stroop colour-word % correct week 3</th>
<th>Stroop colour-word % correct week 4</th>
<th>Stroop colour-word % correct week 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>13</td>
<td>100</td>
<td>85</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>15</td>
<td>36</td>
<td>33</td>
<td>12</td>
<td>46</td>
<td>15</td>
</tr>
<tr>
<td>22</td>
<td>91</td>
<td>85</td>
<td>87</td>
<td>89</td>
<td>15</td>
</tr>
</tbody>
</table>
Figure 7.1. Stroop test results for participant 13.

Figure 7.2. Stroop test results for participant 15.

Figure 7.3. Stroop test results for participant 22.
7.3.2 Trail-making tests

This section presents the trail-making test results for all the participants. Each trail-making test score comprises the number of correctly joined target circles, and a number of incorrect choices. Percentage correct is calculated as

\[
\text{Percentage correct} = \frac{\text{correct}}{\text{correct + incorrect}} \times 100.
\]

Table 7.8 shows the results for trail-making test A, while Table 7.9 shows the results for trail-making test B. Figure 7.4 - Figure 7.6 (overleaf) show the trail-making results for the three participants individually.

Table 7.8. Trail-making (A) results.

<table>
<thead>
<tr>
<th>Participant ID</th>
<th>Trail-making (A) % correct week 1</th>
<th>Trail-making (A) % correct week 2</th>
<th>Trail-making (A) % correct week 3</th>
<th>Trail-making (A) % correct week 4</th>
<th>Trail-making (A) % correct week 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>13</td>
<td>86</td>
<td>89</td>
<td>100</td>
<td>100</td>
<td>96</td>
</tr>
<tr>
<td>15</td>
<td>74</td>
<td>69</td>
<td>66</td>
<td>57</td>
<td>5</td>
</tr>
<tr>
<td>22</td>
<td>96</td>
<td>100</td>
<td>64</td>
<td>81</td>
<td></td>
</tr>
</tbody>
</table>

Table 7.9. Trail-making (B) results.

<table>
<thead>
<tr>
<th>Participant ID</th>
<th>Trail-making (B) % correct week 1</th>
<th>Trail-making (B) % correct week 2</th>
<th>Trail-making (B) % correct week 3</th>
<th>Trail-making (B) % correct week 4</th>
<th>Trail-making (B) % correct week 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>13</td>
<td>96</td>
<td>100</td>
<td>89</td>
<td>83</td>
<td>81</td>
</tr>
<tr>
<td>15</td>
<td>58</td>
<td>60</td>
<td>54</td>
<td>58</td>
<td>57</td>
</tr>
<tr>
<td>22</td>
<td>83</td>
<td>96</td>
<td>83</td>
<td>78</td>
<td></td>
</tr>
</tbody>
</table>
Figure 7.4. Trail-making results for participant 13.

Figure 7.5. Trail-making results for participant 15.

Figure 7.6. Trail-making results for participant 22.
7.3.3 Reaction time test

This section presents the reaction time test results. Reaction time tests were taken three times each session, and the mean is given for each session. Table 7.10 shows the reaction time results for each participant, for each test session. Figure 7.7 - Figure 7.9 (overleaf) show the results for each participant graphically.

<table>
<thead>
<tr>
<th>Participant ID</th>
<th>Mean reaction time week 1</th>
<th>Mean reaction time week 2</th>
<th>Mean reaction time week 3</th>
<th>Mean reaction time week 4</th>
<th>Mean reaction time week 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>13</td>
<td>387</td>
<td>532</td>
<td>353</td>
<td>378</td>
<td>327</td>
</tr>
<tr>
<td>15</td>
<td>2333</td>
<td>1782</td>
<td>2074</td>
<td>1053</td>
<td>698</td>
</tr>
<tr>
<td>22</td>
<td>339</td>
<td>512</td>
<td>349</td>
<td>284</td>
<td></td>
</tr>
</tbody>
</table>
Figure 7.7. Reaction time results for participant 13.

Figure 7.8. Reaction time results for participant 15.

Figure 7.9. Reaction time results for participant 22.
7.3.4 ‘Percentage change towards normal’ analysis

For comparison with the findings of Malec et al. (1984), the ‘percentage change towards normal’ analysis, described in §5.1.2, was undertaken for the measures common to their study and the present study. The calculation performed was:

\[ C_n = \frac{A_n - A_{n-1}}{X - A_{n-1}} \times 100 \]

where

- \( n \) is the test session number: 1 for pre-test, 2 for end of week 1, etc.,
- \( A \) is test score,
- \( X \) is the ‘lower limit of normal’ for the measure, and
- \( C \), the result, is the ‘percentage change towards normal’.

Malec et al. provide values for \( X \) for the Stroop and reaction time tests, but do not provide an \( X \) value for the trail-making tests. To perform the calculation for those measures, a value could be estimated (derived from the scores of the neurotypical participants, perhaps). Alternatively the best possible value could be chosen - in which case the calculation would give the percentage change towards ‘ideal’ rather than ‘normal’. There is an assumption in this analysis that the participants’ scores start off lower than normal, and improve. Using the ‘best possible’ value preserves this assumption. Table 7.11 shows the value of \( X \) used for each measure.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Value for ‘( X )’ (lower limit of normal) for ‘percentage change towards normal’ analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stroop word (number of correct choices in 45 seconds)</td>
<td>88</td>
</tr>
<tr>
<td>Stroop colour</td>
<td>65</td>
</tr>
<tr>
<td>Stroop colour-word</td>
<td>35</td>
</tr>
<tr>
<td>Reaction time</td>
<td>35ms</td>
</tr>
<tr>
<td>Trail-making percent correct</td>
<td>100</td>
</tr>
</tbody>
</table>
The calculated $C$ scores were grouped into three categories, for two criterion levels. The three groups are as follows:

- **Positive:** $C_n$ positive and greater magnitude than the criterion level;
- **Negative:** $C_n$ negative and greater magnitude than the criterion level;
- **Zero:** $C_n$ has lesser magnitude than the criterion level, regardless of sign.

The two criterion levels are 25% and 10%. Statistical analysis was performed using the Wilcoxon ranked sum test for comparison with Malec et al. (1984). This analysis found no significant ($p < .05$) difference in the categorised results between ‘game’ and ‘no game’ weeks, for the Stroop tests and reaction time test, at either the 10% or 25% criteria. The only significant result at the $p < .1$ level was for the reaction time test at the 25% criterion.

**7.3.5 Log in data**

Server logs were analysed to find the typical session duration (time between logging in and logging out or closing the software), and how many players were logged in at any given time. There were 251 distinct login sessions between the start and end dates of the study.

One of the unknown quantities at the start of the study was how long session lengths should be. Figure 7.10 is a histogram of session duration, showing that most sessions were less than 5 minutes long. 33 sessions had durations recorded of more than one hour, up to one day (the maximum allowed): given the distribution of durations, these outliers have been ignored. The most likely causes for these very long session durations are that either the client software crashed, or was left running with no user activity.

![Figure 7.10. Session duration frequency over the study period.](image)
The number of players logged in at the same time varied between 0 and 10, with a median value of 1. In other words, for most of the time, there was only one person online at a time. This is shown in Figure 7.11.

![Number of players logged in concurrently](image)

**Figure 7.11.** Number of players logged in over the duration of the study.

### 7.4 Discussion

This section discusses the results presented above. Of course, the low number of participants must be taken into account. If the results for all the participants exhibit the same trend, this could be taken as an indication of greater external validity, but care must be exercised when extrapolating to a larger group.

As mentioned above, participant 22 was too ill to participate in the final week of the study. Other than that, compliance was excellent, with the participants taking all the tests every week.

Comparing the difference in test scores between ‘game’ and ‘no game’ phases gives an indication of the effect of the intervention. Houle (2009) suggests that positive results should be detectable by visual inspection of a plot of the time-series data collected in a SSRD. Ideally, every test score would improve from a ‘no game’ week to a ‘game’ week; going from a ‘game’ week to ‘no game’ week, the score would stay the same or return to
its baseline value. The results in the present study are mixed. Some of the charts do have a pattern similar to the ideal, while others steadily improve regardless of the intervention. Others charts show the opposite of the ideal pattern, or a gradual worsening of test score. Overall, the results from visual inspection are inconclusive.

Percentage improvement from beginning to end of the study period was calculated, for each participant, for each measure. The percentage improvement from beginning to end of the study did not show a consistent improvement (i.e. an improvement for all three participants), for any measure except reaction time. For participants 13 and 22 the drop is similar, around 15%. For participant 15 the improvement in reaction time is more marked, dropping around 70%. But fluctuations from one week to the next do not follow the ‘ideal’ SSRD pattern. This indicates that the improvements seen are likely to be due to practice effects, and not from playing the multi-player game.

As the present study was based on, and an extension of the study of Malec et al. (1984), the same analysis was performed. The ‘percentage change towards normal’ calculation results in a value which changes from week to week, for each test score, for each participant. The analysis involves categorising these values into three groups, depending on whether the score improved by more than the criterion, worsened by more than the criterion, or was within the same range as the previous score. In the present study, this analysis found that reaction time improved more after a week of game-playing than after a ‘no game’ week, but this improvement was not statistically significant at the $p < .05$ level.

Turning to the other type of data gathered during the study, the server logs were analysed to find the extent of interaction between the players. It was found that for much of the experiment, the participants did not play at the same time. The most important element of the intervention under test was that participants logged in at the same time and played together, in order to gain the proposed benefits of social gaming, and so, the results do not reflect the effect of playing together in a shared, multi-player environment. The need for the players to log in together had been recognised during the study design phase, and addressed by asking participants to log in at an agreed time on ‘game’ days. The server logs have shown that having an agreed-upon time was not sufficient to ensure that all the participants played together in this study.

Session duration was under five minutes in most cases. This is partly explained by the sessions in which no game playing was allowed, with test-taking the only available activity. But the login data includes the neurotypical participants who were invited to play, but did not have to take the tests. One possible explanation for the short average duration is that players logged in, and then quickly logged back out having found that there were no
other players online. Another possible explanation is that the game design, described in §6.3.2, is too simple, and the game is not engaging enough as a result.

Possible reasons for the results being inconclusive are now discussed. With such a low number of participants, wide variation can be expected with few common trends. Within each participant’s individual results, the ‘ideal SSRD’ pattern was seen sporadically, but not consistently for any participant. Recovery from brain injury is unpredictable and uneven, but the time since injury for the participants is greater in the present study than for Malec et al., and should militate against spontaneous improvements. It is possible that the game does not influence the cognitive skills measured by the tests, and that the participants have not responded in the same way to the intervention, giving improvements in different test scores. It may be that the time spent in the game was insufficient to cause improvement, and the low number of participants online concurrently could have contributed to the low engagement.

The health of the individual participants did vary over the course of the study. It may be that some fluctuations from week to week are more due to health or other issues rather than the influence of the multi-player game. Some data points in the results are so anomalous that they cannot be taken at face value. For example, the number of errors in the trail-making (A) test for participant 15 rockets from 20 in week 4 to 80 in week 5. Possible explanations for this type of result are that the participant was interrupted during the test, or perhaps was bored or frustrated.

7.5 Conclusions

In this experiment, the hypothesis was

\[ \textbf{H1:} \text{Multiplayer online video games contribute to an improvement in cognitive function.} \]

This hypothesis would have been supported by an improvement in cognitive test scores following the weeks in which the participants played the multi-player online game. From inspection of the results for each participant, there is not an unambiguous improvement in the test scores at the end of the game weeks. Therefore, the results do not show evidence to support H1 at this stage.

An overall improvement in reaction time was seen between the start and end of the study. This is in agreement with results of previous studies, which have found that playing video games improved reaction time (Green and Bavelier, 2006). However, the game was only played for one-half of the overall experimental schedule, and so it is most likely that much of the improvement in reaction time is due to simple practice effects. The Wilcoxon ranked
Sum test was used to analyse the results. Although the change in reaction time did correlate with whether or not the game was played to some extent, this was not significant at the 5% level.

Improvements were seen in some other test scores (e.g. the Stroop test correct scores), but other measures were worse at the end of the study (e.g. trail-making test incorrect scores). For all measures, the standard deviation in percentage change is large.

Server logs were analysed to find the extent of interaction between the players, as this was the proposed intervention under test. It was found that for much of the experiment, the participants did not play at the same time. The most important element of the intervention under test was that participants logged in at the same time and played together, in order to gain the proposed benefits of social gaming. The results do not reflect the effect of playing together in a shared, multi-player environment. Thus, the main problem found with the experimental design was its reliance on the participants all logging in at the same time.

When the participants are geographically dispersed, it is difficult to ensure this happens. One solution to this problem would be to recruit many more participants, so that there would be more chance of some number of participants online at any given time. Another approach is to modify the game design so that the players have a social experience without needing to be online at the same time.

The next chapter describes the next iteration of the experiment, for which the method and software were revised to reflect these findings.
8 Experiment #2: Measuring wellbeing

This chapter describes the second experiment, conducted over the period May-June 2014. The hypotheses tested in this experiment were:

**H1**: Multiplayer online video games contribute to an improvement in cognitive function.

**H2**: Multiplayer online video games contribute to an improvement in subjective loneliness.

**H3**: Multiplayer online video games contribute to an improvement in subjective quality of life.

In experiment #1, H1 was tested and not proved: the results did not show any convincing improvements in cognitive test scores that could be ascribed to playing the prototype multi-player online game. That raised two questions:

- Do the players simply need to play the game for longer before an improvement in the cognitive test scores can be seen?
- Are we using the right tests to measure improvement?

In the conceptual model for the proposed benefits, it is assumed that cognitive improvements arise in tandem with emotional wellbeing, as it has been proposed by Sohlberg and Mateer (2001) that emotional and cognitive improvements are inseparable. Adding direct measures of quality of life avoids the reliance on this assumption: it may be the case that the cognitive tests are not sensitive to an improvement as measured by the quality of life tests.

The main methodological issue discovered in experiment #1 was that the participants did not often log in at the same time so many of the potential social interactions did not occur. The sense of belonging to a social network is essential to alleviating social loneliness, according to Weiss (1973). It was critical to address this issue in the design of experiment #2, and the design objectives therefore were to:

- Provide more of a sense of community to the players, and
- Ensure that the players experienced more social interactions within the game.

8.1 Changes to the experimental protocol

The duration of each phase (i.e. ‘game’ or ‘no game’ stage of the study) was one week in experiment #1. This was based on the method of Malec et al. (1984) but it was not known whether or not this duration was optimal. As the results from experiment #1 did not show
improvements that could be ascribed to the video game intervention, the duration for each phase was increased from one week to two weeks, giving more time for the proposed beneficial effects to develop. The duration was doubled as the first step in an exponential search (Bentley and Yao, 1976) for the optimal duration, a suitable choice of algorithm where the search space is unbounded.

The measures in the 2013 experiment were broad-based cognitive tests designed to show any improvement in executive function. A therapeutic benefit which would arguably be just as important, if not more so, is an improvement in perceived quality of life. Therefore, two new tests were incorporated into the game software in an attempt to measure this, and test hypotheses H2 and H3. These tests, the De Jong Gierveld Loneliness Scale (De Jong Gierveld and Van Tilburg, 2006) and the Satisfaction With Life Scale (Diener, Emmons, Larsen, and Griffin, 1985), are short questionnaire-style tests (described in §3.3). These tests were integrated into the game software, as for the existing cognitive tests.

In this experiment, the neurotypical participants, as well as the brain-injured participants, were asked to take all the tests each week. This decision was taken for the following reasons:

- The new tests introduced in this experiment were untried, and it was felt that it would be most valuable to collect as much data as possible.
- It may be the case that the new tests detect improvements for both brain-injured and neurotypical participants.

To improve participation and compliance rates, the participants were sent daily emails describing in-game activity. This was intended to incentivise the participants to play the game. The emails were only sent in the second of the two ‘game’ phases of the experiment, allowing a comparison to be made. This daily email is described in more detail in §8.3.

The participants were asked to take the tests again approximately six months after the start of the experiment. The intention of this follow up test was to find out if any effects seen in the results are long lasting. Additionally, participants were asked to answer a very short questionnaire, consisting of three questions, by providing free text responses by email. This qualitative data was intended to provide further illumination of the participants’ experiences and their quantitative results.

Increasing the study length carried the risk of increasing the participant dropout rate. To mitigate this, new consent forms asked the participants to agree to their data being used even if they withdrew from the study. The Faculty of Technology Ethics Committee at the
University of Portsmouth granted ethical approval for these changes to the experimental protocol.

8.2 Changes to the game design

This section details the additions made to the game. There are two main new features. The first is a series of ‘community’ screens, designed to give more of a sense of belonging to a social network of players. The second is a trading feature that provides a turn-based mechanism for interaction to take place. Both these features are intended to tackle the problem that players may only occasionally be online at the same time.

8.2.1 New ‘community’ screens

The new version of the game features new community-based ‘pages’. The first new page is titled ‘what’s been happening’, shown in Figure 8.1 (overleaf). This page appears once per session, after the player has logged in. The purpose of the page is to give the player a sense of being part of a community of players who have been active while the player has been away.

After the player has viewed the ‘what’s been happening’ page, they are taken to their ‘home page’. This screen shows the player’s avatar with their statistics (i.e. health, food and treasure scores). Below this are some community-based options: players can send and receive messages, view other players’ home pages and view their ‘guestbook’. Figure 8.2 (overleaf) shows the home page for the currently logged in player (Jason); clicking the button to see the other players displays the home page for another player (Ella), shown in Figure 8.3 (p. 116). From there, the ‘previous’ and ‘next’ buttons cycle through all the players in the community.
Figure 8.1. ‘What's been happening’ page.

Figure 8.2. Player Jason’s home page.
A player can view his or her own guestbook from their home page, or any other player’s guestbook from that player’s home page. Any player can post messages on any other player’s guestbook. For example, Figure 8.4 shows Ella’s guestbook as seen by player Jason, who can add a new comment to Ella’s guestbook.

Each player can see messages sent and received. Messages are broadcast to all the players when a game event happens, such as losing health. This is so that the other players are alerted to the plight of the player and can help by providing food, which replenishes health.
Figure 8.5 shows messages sent by player Jason, which consist in this example of messages generated by game events: logging in, losing health, and dying.

8.2.2 New trading mechanic

In experiment #1, participants rarely logged in at the same time. For experiment #2, the game software was redesigned to allow for multi-player interaction without logging in at the same time, using a trading game mechanic. This is intended to engage players in a sense of community by providing a game feature that necessitates dealing with other players. Crucially, this game mechanic does not need the players to be online at the same time: to make a trade, one player sends a message to another player, requesting a trade. When the second player logs in, they will see the trade request in their message inbox. To make a trade, the player clicks the button on the home page of the player with whom they wish to trade. This brings up the ‘choose your trade’ screen, shown in Figure 8.6 (overleaf). There are two types of commodity that can be traded: food and treasure. A player who is running low on food can try to trade some of their treasure for food from players with a surplus. On the other hand, the player with surplus food can trade some for treasure collected by other players. This leads to the ‘make a trade’ screen, shown in Figure 8.7 (overleaf), where the player completes the details in a form letter to the other player. The black rectangles in the ‘make a trade’ screen are editable text boxes. When a player receives a trade request, they can opt to go ahead with the trade, refuse, or send another message back, negotiating a new deal.
Figure 8.6. ‘Choose your trade’ screen.

Figure 8.7. ‘Make a trade’ screen.
8.3 The ‘daily digest’ player-incentivising emails

In the second ‘game’ phase of the experimental period, each player was sent a ‘daily digest’ email describing activity in the game. The script that produced each email also generated extra game events, which affected players randomly. The intention behind the emails and the generated activities was to improve player participation in the game. As the emails were sent for only one of the two game periods, some comparison can be made between participation with and without them.

There were three types of procedurally generated game events:

- A player could be told that they feel unwell today, and their health count reduced;
- A player’s food could ‘go off’, reducing their food supply, and
- Some of a player’s treasure could be ‘stolen’, reducing their treasure count.

These occurrences befell any player who did not log in for three days. The intention was that receiving this news would spur the player to re-engage with the game to recover the lost health, food, and treasure.

The email consists of a summary of the player’s status, and notifications of any generated events, followed by up to ten of the most recent in-game happenings. An example email is shown in Figure 8.8 (overleaf). The emails include thumbnail images of the recipient, and other players mentioned.
8.4 The new measures

Two new tests were introduced in this version of the experiment, intended to directly measure the emotional impact of playing a community-based game, testing hypotheses H2 and H3. The tests measure perceived loneliness and satisfaction with life.

8.4.1 De Jong Gierveld Loneliness Scale

The purpose of including this test was to test H2, i.e. to find out if there is a measurable improvement to the perceived isolation of participants, which may be attributable to playing the online multi-player video game. Commonly used loneliness tests were reviewed in §3.3. The chosen scale, the De Jong Gierveld 6-point Loneliness Scale, consists of six items. The answers to these six questions are used to derive scores for social and emotional loneliness, identified by Weiss (1973), and described in §2.2.5. This ability to distinguish between social and emotional loneliness was considered an advantage over
other instruments, as was its brevity, as it was felt important that the tests did not fatigue participants.

In the current software implementation, each question is presented one at a time, with the participant choosing the answer by clicking on a button. The implementation had also to allow for the participant to decline answering. Figure 8.9 shows the loneliness scale as implemented in the current study.

![Figure 8.9. Screenshot showing one example question from the De Jong Gierveld 6-point Loneliness Scale implementation.](image)

For each question, a value is recorded between one (‘Yes!’) and five (‘No!’) inclusive. If a participant declines to answer a question, a value of zero is recorded. The emotional and social loneliness scores are calculated as follows: the emotional loneliness score is the number of positive or neutral answers to the first three questions. The social loneliness score is the number of negative or neutral answers to the final three questions.

### 8.4.2 Satisfaction With Life Scale

To test H3, a test which measures perceived quality of life (QOL) was also added to the tests. The expected result was that perceived QOL would improve in a way that matched the Loneliness Scale results, correlating with the periods of multi-player game activity. The Satisfaction With Life Scale used here was developed by Diener, Emmons, Larsen, and Griffin (1985) and has been used and validated many times since then (Pavot and Diener, 2008). It is another short instrument (5 items), which again is seen as an advantage over longer instruments.

Each question is presented one at a time, with the participant choosing the answer by clicking on a button. The implementation must also allow for the participant to decline
answering. Figure 8.10 and Figure 8.11 shows the Satisfaction With Life Scale as implemented in the present study. In Figure 8.10, the first version is shown, with the standard seven-item scale. The seven choices make the test look complicated, and so were reduced to five, to make the test visually simpler, and easier to understand. This simplified version of the test was validated by Kobau, Sniezek, Zack, Lucas, and Burns (2010). Figure 8.11 shows the second iteration, with the five-item scale.

Figure 8.10. Satisfaction With Life Scale example question showing implementation with seven-item scale.

Figure 8.11. Simpler Satisfaction With Life implementation with five-item scale.
Values recorded are 1 (‘Strongly disagree’) to 5 (‘Strongly agree’) inclusive. If a participant declines to answer a question, a value of zero is recorded. The satisfaction with life score is the sum of these five recorded values, giving a range of (5, 25).

8.5 Method

The method used in experiment #2 was mostly the same as for experiment #1, described in chapter 7. Some changes were made in the light of experience, as described in the sections above, and these are now summarised:

- Phase duration was doubled from one to two weeks,
- Duration of the entire experiment was doubled from four to eight weeks,
- New community and trading features were added to the game,
- Two new measures were added to the suite of tests,
- Daily emails were sent in the final ‘game’ phase of the experiment, and
- Follow up tests and a short questionnaire were administered six months after the experiment.

At the end of the experiment, each participant was sent a ‘thank you’ email with plots of all of his or her own results.

8.5.1 Participant recruitment

In parallel with the development work, recruitment activity continued. Recruitment activities for experiment #2 followed the same method as for experiment #1, i.e. visiting the brain injury support groups Different Strokes and Headway in London. It was hoped that this would bring in another small cohort of brain-injured participants. The recruitment activity did generate interest among potential participants, but unfortunately no new brain-injured participants joined the research project.

The three brain-injured participants from experiment #1 were good enough to volunteer once again. One of these three, participant 22, died in June 2014 following illness. She was able to participate at the start of experiment #2 but only for the first phase.

Additional non-brain injured participants were recruited from a game programming event in January 2014 (‘Global Game Jam’), hosted in London. The mix of non-brain-injured participants changed, but the size of the group did not get any bigger: in total, 16 of the recruited participants logged in to the game and took the tests.

Prior to taking part, all participants were given an information sheet, shown in Appendix 2, and a link to the project website which provided full details of the project. All participants were physically able to sign the consent form.
8.5.2 Schedule

Table 8.1 (overleaf) shows the schedule for the experiment run in 2014. The start date was 9 May 2014, which marked the beginning of a two-week ‘no game’ period. The first ‘game’ period ran from 23 May to 5 June, and the second ‘game’ period from 20 June to 3 July. The last scheduled date of the experiment was 4 July 2014. The experiment was thus split into four phases, the dates of which are shown in Table 8.2 (p. 126). Participants were invited to play the game on any and all days during the ‘game’ phases, and could also take tests on any day (only once per day). The intention of allowing tests on days other than Fridays was to reduce the number of gaps in the data due to missed tests. Additionally, more test data should provide more reliable results, with fewer outliers.

The test scores taken on the first day of a new phase contribute to the mean for the previous phase, not the new phase. The rationale for this is that on the first day of a new phase, the new intervention will not yet have had time to have an effect.
Table 8.1. Schedule for experiment #2.

<table>
<thead>
<tr>
<th>Friday</th>
<th>Saturday</th>
<th>Sunday</th>
<th>Monday</th>
<th>Tuesday</th>
<th>Wednesday</th>
<th>Thursday</th>
</tr>
</thead>
<tbody>
<tr>
<td>9 May 2014</td>
<td>10</td>
<td>11</td>
<td>12</td>
<td>13</td>
<td>14</td>
<td>15</td>
</tr>
<tr>
<td>Take tests</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>17</td>
<td>18</td>
<td>19</td>
<td>20</td>
<td>21</td>
<td>22</td>
</tr>
<tr>
<td>Take tests</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>24</td>
<td>25</td>
<td>26</td>
<td>27</td>
<td>28</td>
<td>29</td>
</tr>
<tr>
<td>Play game and take tests</td>
<td>24</td>
<td>25</td>
<td>26</td>
<td>27</td>
<td>28</td>
<td>29</td>
</tr>
<tr>
<td>30</td>
<td>31</td>
<td></td>
<td>1 June 2014</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Play game and take tests</td>
<td>31</td>
<td>1 June 2014</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
</tr>
<tr>
<td>Take tests</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>14</td>
<td>15</td>
<td>16</td>
<td>17</td>
<td>18</td>
<td>19</td>
</tr>
<tr>
<td>Take tests</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>21</td>
<td>22</td>
<td>23</td>
<td>24</td>
<td>25</td>
<td>26</td>
</tr>
<tr>
<td>Play game and take tests</td>
<td>21</td>
<td>22</td>
<td>23</td>
<td>24</td>
<td>25</td>
<td>26</td>
</tr>
<tr>
<td>27</td>
<td>28</td>
<td>29</td>
<td>30</td>
<td>1 July 2014</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Play game and take tests</td>
<td>28</td>
<td>29</td>
<td>30</td>
<td>1 July 2014</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Take tests</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 8.2. Dates of phases in experiment #2.

<table>
<thead>
<tr>
<th>Phase</th>
<th>Start date (inclusive)</th>
<th>End date (inclusive)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>05/05/14 (and tests taken a day or two early also allowed)</td>
<td>23/05/14</td>
</tr>
<tr>
<td>2</td>
<td>24/05/14</td>
<td>06/06/14</td>
</tr>
<tr>
<td>3</td>
<td>07/06/14</td>
<td>20/06/14</td>
</tr>
<tr>
<td>4</td>
<td>21/06/14</td>
<td>4/7/2014 (and tests taken a day or two late also allowed)</td>
</tr>
</tbody>
</table>

8.6 Results

This section presents the results of the experiment:

- The scores for the new measures, loneliness and satisfaction with life,
- The cognitive test scores (Stroop, trail-making and reaction time tests), and
- The qualitative responses to the three-item questionnaire after six months.

In each phase, a participant could take a test multiple times – the software would allow the tests to be taken up to once per day, although this did not happen in practice. The test schedule stipulated that the tests would be taken twice per phase, and could be taken more often. Thus, the results show the mean of the scores for each phase. The results include the six-month follow-up test scores, for those participants who took them. But before getting to the test results, data showing engagement and compliance are presented.

8.6.1 Engagement and compliance

Some participants did not complete the entire study, or participated sporadically. The effect of these gaps in the data is magnified by the need to have data in two contiguous phases to see the change from one phase to the next. Table 8.3 (overleaf) shows the number of active participants at each phase.
Table 8.3. Number of active participants in each phase.

<table>
<thead>
<tr>
<th></th>
<th>Phase 1</th>
<th>Phase 2</th>
<th>Phase 3</th>
<th>Phase 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of participants who logged in to game</td>
<td>16</td>
<td>14</td>
<td>7</td>
<td>13</td>
</tr>
<tr>
<td>Number of participants who took tests at least once</td>
<td>16</td>
<td>12</td>
<td>7</td>
<td>10</td>
</tr>
</tbody>
</table>

Table 8.4 (overleaf) shows how many times each participant took the tests, their number of logins, and total time spent logged in. Five participants took the tests in every phase, providing a complete series of results. Even so, the number of Fridays (i.e. test days) in the test schedule was nine, and only one participant took the tests the full nine times. As another measure of engagement, Figure 8.12 (p. 129) shows the total time each participant spent logged in to the software. The players who took the tests the most, spent the most time logged in.

There were a total of 157 login sessions over the study period (excluding logins by the researcher, of course). Figure 8.13 (p. 129) shows the number of login sessions for each phase, while Figure 8.14 (p. 130) shows total session duration in each phase. Figure 8.15 (p. 130) shows session length frequency. Sessions lengths of more than one hour are assumed to be due to the software crashing, or the process left running with no user activity.

To show the impact of the new trading game mechanic, appendix 3 shows all the trades offered and accepted. The researcher initiated most trades, in an attempt to prompt the participants to engage with the new game mechanic. Excluding the researcher-initiated trades, nine offers of trades were made, and five were accepted. A very small number of in-game messages were sent. No participants made use of the guestbook feature.
Table 8.4. Number of test sessions per participant.

<table>
<thead>
<tr>
<th>Participant ID</th>
<th>Number of logins</th>
<th>Total time logged in (hours:mins)</th>
<th>Number of times tests taken</th>
<th>Tests taken in all 4 phases?</th>
<th>Follow up tests taken at 6 months?</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>7</td>
<td>9:15</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>5</td>
<td>1:06</td>
<td>3</td>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td>13 (brain-injured)</td>
<td>13</td>
<td>4:40</td>
<td>6</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>15 (brain-injured)</td>
<td>28</td>
<td>10:24</td>
<td>9</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>22 (brain-injured)</td>
<td>5</td>
<td>0:43</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>13</td>
<td>3:48</td>
<td>7</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>7</td>
<td>1:39</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>3</td>
<td>0:27</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>10</td>
<td>3:19</td>
<td>8</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>7</td>
<td>1:24</td>
<td>4</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>16</td>
<td>5:00</td>
<td>8 (but irregularly spaced)</td>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td>36</td>
<td>10</td>
<td>6:12</td>
<td>7</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>38</td>
<td>5</td>
<td>2:04</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>41</td>
<td>6</td>
<td>2:08</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>44</td>
<td>5</td>
<td>3:11</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>5</td>
<td>0:49</td>
<td>2</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 8.12. Total time spent logged in to the system, for each participant.

Figure 8.13. Total number of login sessions, for all 16 participants, in each phase.
8.6.2 Loneliness Scale

The tables in this section show the loneliness scores for each participant, for each phase. Table 8.5 (overleaf) shows the total loneliness scores, Table 8.6 (p. 132) the social loneliness scores, and Table 8.7 (p. 132) the emotional loneliness scores. The emotional loneliness score is the count of neutral and positive answers for questions 1, 2, and 3 (listed in Table 3.2), while the social loneliness score is the count of neutral and negative answers to questions 4, 5, and 6. This gives a score for both in the range (0, 3). Total loneliness is the sum of these two scores, where 0 means ‘least lonely’ and 6 means ‘most lonely’. Gaps in the data are due to participants not taking tests during the phase. (In principle, gaps
could also be due to participants declining to answer, but that was not the case here.)
Within one phase, each participant may have taken the test multiple times, so the tables
show the mean of all scores for each phase. Where a participant took the follow-up tests
(six months from the start of the experiment) those results are also shown. Participants
with an acquired brain injury are marked with ‘ABI’. Figure 8.16 to Figure 8.20 (pp. 133-
134) chart the loneliness scores for each phase, for each participant who took the test in all
four phases of the experiment. Differences between intervention and normal care phases
should be visually apparent when the data is displayed in this way (Houle, 2009). The
expected result would show a reduced loneliness score for phase 2 (first ‘game’ phase)
compared with phase 1 (‘no game’ phase), and similarly a reduction for phase 4 (second
‘game’ phase) compared to phase 3 (second ‘no game’ phase). Scores for the non-
compliant participants are not charted, as with gaps in the data, this visual analysis is not
possible.

Table 8.5. Total loneliness for each participant, for each phase. Brain-injured participants are
marked ‘ABI’.

<table>
<thead>
<tr>
<th>Participant ID</th>
<th>Mean total Loneliness phase 1</th>
<th>Mean total Loneliness phase 2</th>
<th>Mean total Loneliness phase 3</th>
<th>Mean total Loneliness phase 4</th>
<th>6-month follow up</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>1.5</td>
<td>2</td>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>3</td>
<td>3</td>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>(ABI) 13</td>
<td>1</td>
<td>2.5</td>
<td>3</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>(ABI) 15</td>
<td>3.33</td>
<td>3</td>
<td>3</td>
<td>3.67</td>
<td></td>
</tr>
<tr>
<td>(ABI) 22</td>
<td>3.25</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>3.33</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>2</td>
<td>3</td>
<td></td>
<td>1.5</td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>4</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.67</td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>0.5</td>
<td></td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>1</td>
<td>0</td>
<td></td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>36</td>
<td>0.67</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>38</td>
<td>2.5</td>
<td></td>
<td></td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>41</td>
<td>3.5</td>
<td></td>
<td></td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>44</td>
<td>5</td>
<td>4.33</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Table 8.6. Social Loneliness scores.

<table>
<thead>
<tr>
<th>Participant ID</th>
<th>Mean social Loneliness phase 1</th>
<th>Mean social Loneliness phase 2</th>
<th>Mean social Loneliness phase 3</th>
<th>Mean social Loneliness phase 4</th>
<th>6-month follow up</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>0.5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>12</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>(ABI) 13</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>(ABI) 15</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>(ABI) 22</td>
<td>2.75</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>1.67</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>0</td>
<td>1</td>
<td></td>
<td>0.5</td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>2</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>0.67</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>36</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>38</td>
<td>2</td>
<td></td>
<td></td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>41</td>
<td>1.5</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>44</td>
<td>2</td>
<td>1.33</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>1</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

### Table 8.7. Emotional Loneliness scores.

<table>
<thead>
<tr>
<th>Participant ID</th>
<th>Mean emotional Loneliness phase 1</th>
<th>Mean emotional Loneliness phase 2</th>
<th>Mean emotional Loneliness phase 3</th>
<th>Mean emotional Loneliness phase 4</th>
<th>6-month follow up</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>(ABI) 13</td>
<td>1</td>
<td>2.5</td>
<td>3</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>(ABI) 15</td>
<td>0.33</td>
<td>0</td>
<td>0</td>
<td>0.67</td>
<td></td>
</tr>
<tr>
<td>(ABI) 22</td>
<td>0.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>1.67</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>2</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.67</td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>0.5</td>
<td></td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>0.33</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>36</td>
<td>0.67</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>38</td>
<td>0.5</td>
<td></td>
<td></td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>41</td>
<td>2</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>44</td>
<td>3</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>2</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 8.16. Loneliness scores for participant 13 (brain-injured).

Figure 8.17. Loneliness scores for participant 15 (brain-injured).
Figure 8.18. Loneliness scores for participant 23 (not brain-injured).

Figure 8.19. Loneliness scores for participant 27 (not brain-injured).

Figure 8.20. Loneliness scores for participant 36 (not brain-injured).
8.6.3 Satisfaction With Life Scale (SWLS)

This section presents the Satisfaction With Life Scale scores. SWLS scores are integers in the range (5, 25), and are the sum of the values of the responses to the five questions listed in Table 3.3. Table 8.8 shows the SWLS score for each participant, for each phase. Within one phase, each participant may have taken the test multiple times, so the table shows the mean of the scores for each phase. Figure 8.21 - Figure 8.25 (pp. 136-137) show the SWLS results for the five participants who took tests in all four phases. The expected result should show an improvement (increase) in SWLS score in ‘game’ phases compared to ‘no game’ phases. As for loneliness, scores for the non-compliant participants are not shown in chart form.

Table 8.8. Mean Satisfaction With Life scores for each participant, for each phase.

<table>
<thead>
<tr>
<th>Participant ID</th>
<th>Mean SWLS phase 1</th>
<th>Mean SWLS phase 2</th>
<th>Mean SWLS phase 3</th>
<th>Mean SWLS phase 4</th>
<th>6-month follow up</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>17.5</td>
<td>18</td>
<td></td>
<td>18.33</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>22</td>
<td>25</td>
<td></td>
<td></td>
<td>25</td>
</tr>
<tr>
<td>(ABI) 13</td>
<td>17.5</td>
<td>20.5</td>
<td>20</td>
<td>19</td>
<td>20</td>
</tr>
<tr>
<td>(ABI) 15</td>
<td>7</td>
<td>9.5</td>
<td>10</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>(ABI) 22</td>
<td>11.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>17</td>
<td>19</td>
<td>18</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>17</td>
<td>13</td>
<td></td>
<td>12.5</td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>15.5</td>
<td></td>
<td>14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>24</td>
<td>23.5</td>
<td>24</td>
<td>24</td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>17.5</td>
<td></td>
<td>18</td>
<td>17</td>
<td>19</td>
</tr>
<tr>
<td>35</td>
<td>21.67</td>
<td>22</td>
<td></td>
<td>24</td>
<td>23</td>
</tr>
<tr>
<td>36</td>
<td>15.67</td>
<td>18</td>
<td>17</td>
<td>16.5</td>
<td></td>
</tr>
<tr>
<td>38</td>
<td>17.5</td>
<td></td>
<td></td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>41</td>
<td>16.5</td>
<td></td>
<td></td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>44</td>
<td>8</td>
<td>13</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>13</td>
<td>16</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 8.21. SWLS results for participant 13 (brain-injured).

Figure 8.22. SWLS results for participant 15 (brain-injured).
Figure 8.23. SWLS results for participant 23 (not brain-injured).

Figure 8.24. SWLS results for participant 27 (not brain-injured).

Figure 8.25. SWLS results for participant 36 (not brain-injured).
8.6.4 Stroop tests

This section presents the Stroop test results for all the participants. Each test yields a number of correct and incorrect choices. The percentage of correct choices is calculated as

$$\text{Percentage correct} = \frac{\text{correct}}{\text{correct} + \text{incorrect}} \times 100.$$ 

Table 8.9 shows the Stroop word percentage correct score for each participant, for each phase. Within one phase, each participant may have taken the test multiple times, so the table shows the mean of the scores for each phase. Table 8.10 (overleaf) shows the Stroop colour test results, and Table 8.11 (overleaf) the Stroop colour-word results. Figure 8.26 - Figure 8.30 (pp. 140-141) show the Stroop test results for the participants who took the tests in all four phases.

Table 8.9. Stroop word results.

<table>
<thead>
<tr>
<th>Participant ID</th>
<th>Stroop word mean % correct phase 1</th>
<th>Stroop word mean % correct phase 2</th>
<th>Stroop word mean % correct phase 3</th>
<th>Stroop word mean % correct phase 4</th>
<th>6-month follow up</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>93.64</td>
<td>97.56</td>
<td>92.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>94.8</td>
<td>95.74</td>
<td></td>
<td>94.45</td>
<td></td>
</tr>
<tr>
<td>(ABI) 13</td>
<td>100</td>
<td>96.875</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>(ABI) 15</td>
<td>92.43</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>(ABI) 22</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>99.12</td>
<td>100</td>
<td>97.14</td>
<td>98.65</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>95.65</td>
<td>93.02</td>
<td></td>
<td>93.65</td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>97.78</td>
<td>97.87</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>100</td>
<td>97.37</td>
<td>100</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>98.65</td>
<td>93.99</td>
<td></td>
<td>97.67</td>
<td>95.65</td>
</tr>
<tr>
<td>35</td>
<td>98.65</td>
<td></td>
<td></td>
<td>91.49</td>
<td>95.56</td>
</tr>
<tr>
<td>36</td>
<td>99.36</td>
<td>96</td>
<td>100</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>38</td>
<td>98.81</td>
<td></td>
<td></td>
<td>97.83</td>
<td></td>
</tr>
<tr>
<td>41</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>44</td>
<td>100</td>
<td>97.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>97.14</td>
<td>97.62</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Participant ID</td>
<td>Stroop colour mean % correct phase 1</td>
<td>Stroop colour mean % correct phase 2</td>
<td>Stroop colour mean % correct phase 3</td>
<td>Stroop colour mean % correct phase 4</td>
<td>6-month follow up</td>
</tr>
<tr>
<td>----------------</td>
<td>-------------------------------------</td>
<td>-------------------------------------</td>
<td>-------------------------------------</td>
<td>-------------------------------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>7</td>
<td>93.32</td>
<td>97.92</td>
<td></td>
<td>98.58</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>97.28</td>
<td>100</td>
<td></td>
<td></td>
<td>98.1</td>
</tr>
<tr>
<td>(ABI) 13</td>
<td>100</td>
<td>98.48</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>(ABI) 15</td>
<td>97.62</td>
<td>96.30</td>
<td>100</td>
<td>95.06</td>
<td></td>
</tr>
<tr>
<td>(ABI) 22</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>100</td>
<td>97.22</td>
<td>100</td>
<td>97.67</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>100</td>
<td>98</td>
<td></td>
<td>96.15</td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>99.12</td>
<td>96.30</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>98.71</td>
<td>100</td>
<td>100</td>
<td>99.19</td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>100</td>
<td>100</td>
<td>97.44</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>98.76</td>
<td>99.07</td>
<td></td>
<td>98.04</td>
<td>100</td>
</tr>
<tr>
<td>36</td>
<td>97.68</td>
<td>98.33</td>
<td>98.25</td>
<td>97.35</td>
<td></td>
</tr>
<tr>
<td>38</td>
<td>100</td>
<td></td>
<td></td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>41</td>
<td>97.73</td>
<td></td>
<td></td>
<td>97.83</td>
<td></td>
</tr>
<tr>
<td>44</td>
<td>100</td>
<td>99.26</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>100</td>
<td>97.73</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Participant ID</th>
<th>Stroop colour-word mean % correct phase 1</th>
<th>Stroop colour-word mean % correct phase 2</th>
<th>Stroop colour-word mean % correct phase 3</th>
<th>Stroop colour-word mean % correct phase 4</th>
<th>6-month follow up</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>95.52</td>
<td>93.18</td>
<td>98.37</td>
<td></td>
<td>93.62</td>
</tr>
<tr>
<td>12</td>
<td>94.54</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(ABI) 13</td>
<td>100</td>
<td>98.21</td>
<td>25</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>(ABI) 15</td>
<td>78.52</td>
<td>90.70</td>
<td>88.89</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>(ABI) 22</td>
<td>89.81</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>98.15</td>
<td>100</td>
<td>100</td>
<td>97.14</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>92.31</td>
<td>100</td>
<td></td>
<td>88.98</td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>93.48</td>
<td>95</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>98.65</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>35</td>
<td>100</td>
<td>94.43</td>
<td>95.74</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>36</td>
<td>98</td>
<td>100</td>
<td>94.44</td>
<td>98.21</td>
<td></td>
</tr>
<tr>
<td>38</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>41</td>
<td>97.73</td>
<td></td>
<td></td>
<td>97.78</td>
<td></td>
</tr>
<tr>
<td>44</td>
<td>97.06</td>
<td>98.48</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>100</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 8.26. Stroop test results for participant 13 (brain-injured).

Figure 8.27. Stroop test results for participant 15 (brain-injured).
Figure 8.28. Stroop test results for participant 23 (not brain-injured).

Figure 8.29. Stroop test results for participant 27 (not brain-injured).

Figure 8.30. Stroop test results for participant 36 (not brain-injured).
8.6.5 Trail-making tests

This section presents the trail-making test results for all the participants. Each trail-making test score comprises the number of correctly joined target circles, and a number of incorrect choices. Percentage correct is calculated as

\[
\text{Percentage correct} = \frac{\text{correct}}{\text{correct} + \text{incorrect}} \times 100.
\]

Table 8.12 (overleaf) shows the results for the trail-making ‘A’ test, while Table 8.13 (overleaf) shows the results for the trail-making ‘B’ test. Figure 8.31 - Figure 8.35 (pp. 144-145) show the trail-making test scores for the participants who took the tests in all four phases.
Table 8.12. Trail-making (A) test results.

<table>
<thead>
<tr>
<th>Participant ID</th>
<th>Trail Making A mean % correct phase 1</th>
<th>Trail Making A mean % correct phase 2</th>
<th>Trail Making A mean % correct phase 3</th>
<th>Trail Making A mean % correct phase 4</th>
<th>6-month follow up</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>84.97</td>
<td>96.15</td>
<td>90.98</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>91.18</td>
<td></td>
<td></td>
<td></td>
<td>83.3</td>
</tr>
<tr>
<td>(ABI) 13</td>
<td>86.31</td>
<td>94.64</td>
<td>89.29</td>
<td>92.59</td>
<td>100</td>
</tr>
<tr>
<td>(ABI) 15</td>
<td>88.96</td>
<td>92.59</td>
<td>86.21</td>
<td>92.68</td>
<td></td>
</tr>
<tr>
<td>(ABI) 22</td>
<td>93.78</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>92.84</td>
<td>100</td>
<td>96.15</td>
<td>92.59</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>96.15</td>
<td>92.59</td>
<td></td>
<td>96.30</td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>93.10</td>
<td>83.33</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>100</td>
<td>98.08</td>
<td>100</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>98.08</td>
<td></td>
<td>92.59</td>
<td>92.59</td>
<td>89.29</td>
</tr>
<tr>
<td>35</td>
<td>100</td>
<td>88.97</td>
<td></td>
<td>89.29</td>
<td>89.29</td>
</tr>
<tr>
<td>36</td>
<td>96.25</td>
<td>100</td>
<td>96.15</td>
<td>57.90</td>
<td></td>
</tr>
<tr>
<td>38</td>
<td>92.72</td>
<td></td>
<td></td>
<td>86.21</td>
<td></td>
</tr>
<tr>
<td>41</td>
<td>96.15</td>
<td></td>
<td></td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>44</td>
<td>100</td>
<td></td>
<td>61.44</td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>83.33</td>
<td>92.59</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 8.13. Trail-making (B) test results.

<table>
<thead>
<tr>
<th>Participant ID</th>
<th>Trail Making B mean % correct phase 1</th>
<th>Trail Making B mean % correct phase 2</th>
<th>Trail Making B mean % correct phase 3</th>
<th>Trail Making B mean % correct phase 4</th>
<th>6-month follow up</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>94.37</td>
<td>80.65</td>
<td></td>
<td></td>
<td>87.07</td>
</tr>
<tr>
<td>12</td>
<td>81.41</td>
<td></td>
<td></td>
<td></td>
<td>71.43</td>
</tr>
<tr>
<td>(ABI) 13</td>
<td>77.39</td>
<td>89.40</td>
<td>96.15</td>
<td>83.33</td>
<td>100</td>
</tr>
<tr>
<td>(ABI) 15</td>
<td>49.81</td>
<td>80.73</td>
<td>69.44</td>
<td>70.44</td>
<td></td>
</tr>
<tr>
<td>(ABI) 22</td>
<td>81.77</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>94.64</td>
<td>89.29</td>
<td>96.15</td>
<td>96.30</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>92.59</td>
<td>89.29</td>
<td></td>
<td>90.94</td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>96.30</td>
<td>80.65</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>100</td>
<td>96.15</td>
<td>100</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>98.08</td>
<td></td>
<td>96.15</td>
<td>89.29</td>
<td>92.6</td>
</tr>
<tr>
<td>35</td>
<td>97.53</td>
<td>90.23</td>
<td></td>
<td>64.10</td>
<td>92.6</td>
</tr>
<tr>
<td>36</td>
<td>97.44</td>
<td>100</td>
<td>89.29</td>
<td>98.08</td>
<td></td>
</tr>
<tr>
<td>38</td>
<td>81.99</td>
<td></td>
<td></td>
<td>92.60</td>
<td></td>
</tr>
<tr>
<td>41</td>
<td>86.76</td>
<td></td>
<td></td>
<td>86.21</td>
<td></td>
</tr>
<tr>
<td>44</td>
<td>100</td>
<td>93.10</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>86.21</td>
<td>86.21</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 8.31. Trail making results for participant 13 (brain-injured).

Figure 8.32. Trail making results for participant 15 (brain-injured).
Figure 8.33. Trail making test results for participant 23 (not brain-injured).

Figure 8.34. Trail making results for participant 27 (not brain-injured).

Figure 8.35. Trail making results for participant 36 (not brain-injured).
8.6.6 Reaction time

Table 8.14 shows the results for the reaction time test for all participants, with the brain-injured participants indicated. Figure 8.36 - Figure 8.40 (pp. 147-148) show the reaction times for the participants who took the test in all four phases of the experiment.

Table 8.14. Reaction time test results.

<table>
<thead>
<tr>
<th>Participant ID</th>
<th>Mean Reaction Time phase 1</th>
<th>Mean Reaction Time phase 2</th>
<th>Mean Reaction Time phase 3</th>
<th>Mean Reaction Time phase 4</th>
<th>6-month follow up</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>243.67</td>
<td>209.33</td>
<td>230.22</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>191.67</td>
<td>295.33</td>
<td></td>
<td>222.67</td>
<td></td>
</tr>
<tr>
<td>(ABI) 13</td>
<td>373</td>
<td>322.5</td>
<td>336.33</td>
<td>268.67</td>
<td>334</td>
</tr>
<tr>
<td>(ABI) 15</td>
<td>286.44</td>
<td>262.17</td>
<td>258.67</td>
<td>281</td>
<td></td>
</tr>
<tr>
<td>(ABI) 22</td>
<td>323</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>281.78</td>
<td>316</td>
<td>272.33</td>
<td>246.83</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>211</td>
<td>217.67</td>
<td></td>
<td>229.83</td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>233.5</td>
<td>232.33</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>442</td>
<td>490</td>
<td>479.67</td>
<td>404.11</td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>294.5</td>
<td></td>
<td>316.67</td>
<td>344</td>
<td>318</td>
</tr>
<tr>
<td>35</td>
<td>180.89</td>
<td>194</td>
<td></td>
<td>200.667</td>
<td>211</td>
</tr>
<tr>
<td>36</td>
<td>215.33</td>
<td>209.33</td>
<td>208</td>
<td>232.67</td>
<td></td>
</tr>
<tr>
<td>38</td>
<td>286.33</td>
<td></td>
<td>324.33</td>
<td></td>
<td></td>
</tr>
<tr>
<td>41</td>
<td>289.5</td>
<td></td>
<td>256.33</td>
<td></td>
<td></td>
</tr>
<tr>
<td>44</td>
<td>354.67</td>
<td>329.78</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>249.67</td>
<td>205.33</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 8.36. Reaction time for participant 13 (brain-injured).

Figure 8.37. Reaction time for participant 15 (brain-injured).
Figure 8.38. Reaction time for participant 23 (not brain-injured).

Figure 8.39. Reaction time for participant 27 (not brain-injured).

Figure 8.40. Reaction time for participant 36 (not brain-injured).
8.6.7 ‘Percentage change towards normal’ analysis

For comparison with the results of experiment #1 and the findings of Malec et al. (1984), the ‘percentage change towards normal’ analysis, described in chapter 5, was undertaken for the measures common to their study and the present study. The calculation performed was:

\[ C_n = \frac{A_n - A_{n-1}}{X - A_{n-1}} \times 100 \]

where

- \( n \) is the phase number,
- \( A \) is test score,
- \( X \) is the ‘lower limit of normal’ for the measure, and
- \( C \), the result, is the ‘percentage change towards normal’.

Values for \( X \) are as follows. The Stroop and trail-making scores are given as percentage correct, so a percentage correct value of 100 is used for \( X \). For SWLS and Loneliness, ‘best possible’ \( X \) values are used. For the SWLS implemented here, the best possible score is 25. For the Loneliness test, the best score is zero. For reaction time, the value from Malec et al. is used. The values for \( X \), then, are as shown in Table 8.15.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Value for ‘( X ) for ‘percentage change towards normal’ analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stroop test percent correct</td>
<td>100</td>
</tr>
<tr>
<td>Reaction time</td>
<td>35ms</td>
</tr>
<tr>
<td>Trail-making percent correct</td>
<td>100</td>
</tr>
<tr>
<td>SWLS</td>
<td>25</td>
</tr>
<tr>
<td>Loneliness</td>
<td>0</td>
</tr>
</tbody>
</table>

The percentage change scores were grouped into three categories, for two criterion levels, 25% and 10%. The three groups are:
• Positive: $C_n$ positive and greater magnitude than the criterion level;
• Negative: $C_n$ negative and greater magnitude than the criterion level;
• Zero: $C_n$ has lesser magnitude than the criterion level, regardless of sign.

The Wilcoxon ranked sum test was used to analyse the categorised scores, only for the participants who took the tests in all four phases. (This analysis would not be meaningful for the participants with gaps in their data.) This showed:

• An improvement between ‘game’ and ‘no game’ phases for total loneliness, at both the 10% and 25% criteria, but these were significant at only $p < .1$.
• The Stroop colour-word test showed a significant improvement between ‘game’ and ‘no game’ phases at the $p < .05$ level, at both the 10% and 25% criteria.
• The Stroop colour results showed the opposite: a significant ($p < .05$) worsening of results in ‘game’ phases.

This last result is surprising as the Stroop colour test is probably the simplest of the three Stroop tests. Perhaps it is too easy, and bores participants, leading to poor performance.

8.6.8 Qualitative results

After the experiment, the participants were asked to comment on their experiences over the study. The researcher emailed each participant with three questions:

1. Did you feel that playing the game gave you a sense of belonging to a community?
2. Do you feel that playing the game had an effect on your sense of loneliness?
3. Do you feel that playing the game had an effect on your sense of satisfaction with life?

Participants 12, 27, 32, 35 and 36 replied. The responses to question 1 are listed in Table 8.16 (overleaf), to question 2 in Table 8.17 (overleaf), and to question 3 in Table 8.18 (p. 152).
Table 8.16. Qualitative responses to the question ‘Did you feel that playing the game gave you a sense of belonging to a community?’

<table>
<thead>
<tr>
<th>Participant ID</th>
<th>Responses to the question ‘Did you feel that playing the game gave you a sense of belonging to a community?’</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>During the time I was able to play the game I felt that my contributions were able to help the other players, therefore giving me a sense of belonging to the community.</td>
</tr>
<tr>
<td>27</td>
<td>It was fun playing the game with my classmates when I was at [college attended by participant], as we figured out the puzzles together, and had a lot of fun.</td>
</tr>
<tr>
<td>32</td>
<td>I felt a sense of community when a couple of players were online at the same time and interacting with each other.</td>
</tr>
<tr>
<td>35</td>
<td>I felt that knowing there were other players participating at the same time as myself and accomplishing the same goals made it feel like a community game, especially with the updates that told me what everyone was accomplishing and the ability to send messages. The lack of interaction during game play (there was no one else on the map with me) made it feel more isolated.</td>
</tr>
<tr>
<td>36</td>
<td>No. I didn't interact with any of the other players.</td>
</tr>
</tbody>
</table>

Table 8.17. Responses to the question ‘Do you feel that playing the game had an effect on your sense of loneliness?’

<table>
<thead>
<tr>
<th>Participant ID</th>
<th>Responses to the question ‘Do you feel that playing the game had an effect on your sense of loneliness?’</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>The email updates on what was going on in game really helped me keep up to date with what everyone was doing when I wasn't able to login. Being able to interact with others, knowing that there are other people was enough to combat my sense of loneliness.</td>
</tr>
<tr>
<td>27</td>
<td>I do not have a sense of loneliness, so I cannot really answer.</td>
</tr>
<tr>
<td>32</td>
<td>No effect - I did not feel lonely nor &quot;popular&quot; (not sure what the opposite to lonely is) when playing the game.</td>
</tr>
<tr>
<td>35</td>
<td>I don't think it did. I would spend an average of 10-15 mins a day on it, compared to the many hours I spend with students and colleagues at work.</td>
</tr>
<tr>
<td>36</td>
<td>No. I didn't get emotionally or psychologically involved enough in the game.</td>
</tr>
</tbody>
</table>
Table 8.18. Responses to the question ‘Do you feel that playing the game had an effect on your sense of satisfaction with life?’

<table>
<thead>
<tr>
<th>Participant ID</th>
<th>Responses to the question ‘Do you feel that playing the game had an effect on your sense of satisfaction with life?’</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>Playing the game and socialising with the other players definitely contributed to my satisfaction with life.</td>
</tr>
<tr>
<td>27</td>
<td>While playing the game was fun, I don't think it had an effect on my satisfaction of life.</td>
</tr>
<tr>
<td>32</td>
<td>I was glad to help out with the research but I don't feel the game gave me any satisfaction with life.</td>
</tr>
<tr>
<td>35</td>
<td>I'm not sure if I can confidently say that. I finished exploring the map after a few days, and remember asking some of the other participants if they had discovered more rooms (their characters appeared to have picked up more diamonds than I did). After that, I'm not sure if the game had any emotional impact on me.</td>
</tr>
<tr>
<td>36</td>
<td>A little as the questions asked were not ones I asked myself from day to day. However, the impact was minimal.</td>
</tr>
</tbody>
</table>

### 8.7 Discussion

This section discusses the data presented above, starting with the login statistics and the usage of the new community and trading features. The results of the cognitive tests are discussed, and the results from the new loneliness and satisfaction with life measures. The follow-up data is compared with predictions made and the qualitative responses are discussed with reference to the corresponding quantitative data.

#### 8.7.1 Engagement

Firstly, engagement in experiment #2 is compared with that of experiment #1. The duration of experiment #2 was twice that of experiment #1, with the same number of active participants, but the number of login sessions in experiment #2 was fewer than in experiment #1. On the plus side, session lengths were longer, but overall less time was spent in the game in experiment #2. Some, but not all, of the participants had also participated in experiment #1, and one reason for the decreased engagement could be that the novelty of the game had worn off.

On the other hand, it could be that quality of interaction is more important than quantity, so time spent in the game may not be the whole story. The doubled study duration potentially gave participants the feeling of belonging to a community for twice as long. MMO players
can feel part of a social network without necessarily interacting with others (Ducheneaut, Yee, Nickell, and Moore, 2006).

During the second game phase, the participants received daily emails intended to improve engagement. Did this intervention have any noticeable impact on engagement? In the qualitative responses, the emails are mentioned once:

‘The email updates on what was going on in game really helped me keep up to date with what everyone was doing when I wasn't able to login.’

The number of participants who logged into the game at phase 4 (second ‘game’ phase) was 13, only one less than the number who logged in during phase 2 (first ‘game’ phase). Total session duration was 15.5 hours in phase 4, down from 24 hours in phase 2. These figures show that engagement dropped from the first to the second half of the experiment, in that the time spent logged in to the game was less. But in terms of retention of the number of players logging in, losing only one player is considered a very good result. With the addition of the qualitative response, the interpretation is drawn that the emails did help to buoy up engagement when interest may have been starting to flag. An unanswered question is: were the procedurally generated game events (reducing players’ health, food and treasure) an important factor in retaining players? Or would simple reminder emails have been just as effective? An ‘A/B’ test in a future iteration of the experiment could answer this.

8.7.2 Loneliness

This section discusses the loneliness results for the five participants who took the loneliness test in all four phases of the experiment. Two of these participants are brain injured. Attempting to interpret the data from the non-compliant participants is not done here as data for ‘game’ and ‘no game’ phases cannot be compared.

Participant 13 is one of the brain-injured participants included in this discussion. His results showed a score of zero for social loneliness over the duration of the experiment. Emotional loneliness worsened from phase 1 (first ‘no game’ phase) to phase 2 (first ‘game’ phase), against expectations. Emotional loneliness worsened again from phase 2 to phase 3, but then improved from phase 3 (second ‘no game’ phase) to phase 4 (second ‘game’ phase).

Participant 15, the second brain-injured participant, showed mixed results for emotional loneliness, with an improvement seen from phase 1 to phase 2, but a worsening from phase 3 to phase 4. It can certainly be seen that participants 13 and 15 have very different
‘loneliness profiles’, with participant 15 having the highest possible score for social loneliness, which did not change over the course of the experiment.

The results for participant 23 (not brain-injured) show an improvement in social loneliness score for the ‘game’ phases compared to ‘no game’ phases, in line with expectations and supporting hypothesis H2. For emotional loneliness, a drop is seen from phase 1 (first ‘no game’ phase) to phase 2 (first ‘game’ phase), but the same pattern is not seen going from phase 3 to phase 4.

Participant 27 recorded a score of zero for social loneliness over the entire study. Emotional loneliness did not change from phase 1 to phase 2, but did improve from phase 3 to phase 4.

Participant 36 also had a score of zero for social loneliness, which did not change over the course of the experiment. For emotional loneliness, an improvement was seen between phases 1 and 2, but this then remained at zero for phases 3 and 4.

Overall, visual inspection of the results for total loneliness yielded only one participant fully matching the prediction of hypothesis H2, while the four others showed partial support, with some contradictory features. This reading is in line with the results of the Wilcoxon ranked sum analysis (§8.6.7), which found some support for H2 but only at $p < .1$ significance.

A point to note is that three participants scored zero for social loneliness, for both ‘game’ and ‘no game’ phases. This could simply mean that those participants did not feel any effects of social loneliness. Another possibility is that the six-item instrument chosen is not sufficiently sensitive. Conversely, one participant recorded maximum values for social loneliness over the entire study. Using a more detailed questionnaire in a future study could reduce these possible floor and ceiling effects, but would have to be balanced with other considerations such as participant boredom and overall time spent taking the tests.

Taking the results as a whole, it was noticed that the two participants with the highest loneliness scores at the start of the study dropped out after the second phase. Could it be the case that loneliness score at the start of the study predicted how many times a participant would take the tests? Figure 8.41 (overleaf) shows the relationship between initial loneliness score for a participant and the number of times that participant took the tests over the course of the study. The correlation coefficient $r$ is -0.27. The negative correlation makes sense as a higher loneliness score means ‘more lonely’, but the magnitude is weak: there are other factors at work that determine how many times a participant takes the tests.
Figure 8.41. Correlation between initial loneliness score and study compliance. (Higher loneliness score means ‘more lonely’.) \( r = -0.27 \).

8.7.3 Satisfaction With Life

This section discusses the Satisfaction With Life Scale (SWLS) results for the five compliant participants. Two of these participants are brain-injured. The first of these two, participant 13, showed improvement from phase 1 (first ‘no game’ phase) to phase 2 (first ‘game’ phase). This was the expected result, but this was not repeated from phase 3 (second ‘no game’ phase) to phase 4 (second ‘game’ phase), which showed instead a drop in SWLS score. The other brain-injured participant, participant 15, did show improvement in SWLS score from phase 1 to phase 2, and again from phase 3 to phase 4, which agrees with the hypothesis, H3.

The results for participants 23 and 36 look similar in shape to those for participant 13, i.e. an improvement from phase 1 to phase 2 that was not repeated in the second half of the study.

Participant 27 was the only one of the five compliant participants to show a drop in SWLS score from phase 1 to phase 2.

Overall then, the results from the first half of the study look promising, but this was not repeated in the second half for most of the participants. This could perhaps point to boredom setting in with the game, or with the study itself, by the halfway point.

A final observation is that the scores for each participant do lie within a fairly narrow range of values, though of course the range for each participant is different. This would seem to agree with the notion that people have a ‘natural’ mean level of satisfaction with life to which they return over time (Myers and Diener, 1995).
8.7.4 The cognitive tests

This section discusses the Stroop, trail-making and reaction time test results for the five compliant participants.

For the Stroop colour-word tests, participants 15 (brain injured) and 36 (not brain injured) showed the ‘ideal SSRD’ shape in their results, i.e. an improvement from phase 1 to phase 2, and an improvement from phase 3 to phase 4. Results for the other participants were more mixed, but overall results were in line with expectations for hypothesis H1. This was borne out by Wilcoxon ranked sum analysis, which did show a significant ($p < .05$) improvement between ‘game’ and ‘no game’ phases. Unfortunately this result was negated by the Stroop colour test results, which showed the opposite! Visual inspection of the Stroop colour results does indeed show results mostly contradicting H1. This is interesting because the Stroop colour-word test should be the more difficult test, and so these findings may be due to participant boredom, and perhaps frustration or fatigue. Finally, visual inspection of the Stroop word test results gives a mixed picture, rather than agreement with either of the other two results. Overall, the Stroop tests do not provide conclusive evidence for support of H1.

The trail-making and reaction time test results show mixed, inconclusive results, with no ‘ideal SSRD’ pattern clearly seen for any participant. As described in §8.6.7, the Wilcoxon ranked sum test was used to analyse the results, using the method given by Malec et al. (1984). Most results did not show a consistent improvement following ‘game’ phases. Overall, conclusive evidence remains to be found for cognitive improvement that can be ascribed to playing the prototype game.

8.7.5 Practice effects and the Hawthorne effect

As for experiment #1 in §7.4, steadily improving results in the cognitive tests could indicate the influence of practice effects. Given the fairly poor take-up of interactions between players, the results should be inspected for signs of improvement that happened regardless of the presence of the intervention. In the case of the cognitive tests, few of the individual results show a continued improvement over the course of the experiment. The reaction time results do display a downward trend for some of the participants, with no significant difference between intervention and normal care phases. This being the case, the slight improvement in reaction time seen for some participants is viewed as evidence of a practice effect.

For the tests of emotional wellbeing, do practice effects apply? Perhaps it is appropriate to consider the Hawthorne effect, which refers to a participant’s awareness of being studied,
and its potential to influence behaviour, skewing results. (The term originated in studies undertaken between 1924 and 1933, at the Western Electric telephone manufacturing factory at Hawthorne, near Chicago.) McCambridge, Witton, and Elbourne (2014) caution that little is known about the mechanisms or magnitude of the Hawthorne effect. Furthermore,

‘In all likelihood, the most common use of the Hawthorne effect term is as a post hoc interpretation of unexpected study findings, particularly where they are disappointing, for example, when there are null findings in trials.’

This suggests that the term should be used with some restraint, but in their systematic review of 19 studies within health sciences, McCambridge et al. did find that ‘consequences of research participation for behaviours being investigated do exist’.

Improvements in loneliness and Satisfaction With Life results seen over the course of the study could therefore be due to participation, and this is probably the case, given the limited levels of social interactions fostered in the prototype experimental game.

### 8.7.6 Follow-up test results

Four participants took follow-up tests after six months, allowing a comparison between their test scores during the study period and six months later. Given the findings so far, some predictions could be made. For the cognitive tests (i.e. Stroop, trail-making and reaction time tests), little difference was found from the start to the end of the study, and no notable improvements were seen that could be ascribed to periods of playing the game. Thus, the follow-up results for these tests were predicted to be close to the mean of the results over the main study period.

Myers and Diener (1995) report that a person’s subjective quality of life will return to its natural level, even after a major life event such as ABI. From this, we can predict that SWLS scores would return to their baseline values after six months. For the loneliness scores, a return to baseline could also apply. These conjectures are of course subject to fluctuations and trends due to other unknown events.

In actual fact the six-month results for the cognitive test results were on track with expectations, but loneliness and SWLS scores were better than predicted. Further exploration of this in a larger study should take into account the rate of attrition: only a quarter of the participants took the follow-up tests. Perhaps only happy and non-lonely participants felt inclined to respond.
8.7.7 Qualitative responses

This section discusses the qualitative responses gathered after six months. Five of the participants replied to the follow-up email questionnaire. None of these participants are brain-injured, and so these responses do not reflect the opinions of the group of most interest to the research. Nevertheless, these replies are presented here as they could provide some insight into the effectiveness of the game features and the effects of playing the game on loneliness and satisfaction with life.

The replies to the first question, ‘Did you feel that playing the game gave you a sense of belonging to a community?’ were mostly positive, especially when other players were online at the same time. Helping other players was mentioned as a positive aspect that enhanced the sense of community, as was receiving updates on other players’ accomplishments. These updates were provided through more than one mechanism – the email updates, the ‘what’s been happening’ page, and the broadcast messaging feature all provided updates on game events. Further research could perhaps find the most effective channels for updates, or simply aim to provide updates in as many forms as possible, given that they are reported as a positive feature.

Regarding questions two and three, most of the respondents felt that playing the game did not contribute to an improvement in satisfaction with life, nor to an improvement in their perception of loneliness.

Of the five participants who responded, participants 27 and 36 took the tests in all four phases, allowing a comparison of their quantitative scores and qualitative opinions. For the other participants, this comparison would be less illuminating due to gaps in the data (compliance for these participants is shown in Table 8.19, overleaf).

For total loneliness, participant 27 scored a consistently low value over the duration of the study. This was comprised of emotional loneliness – for social loneliness, this participant scored zero throughout the study (this is the lowest possible value, meaning ‘not lonely’). Turning to satisfaction with life, for participant 27, SWLS started at a high score of 24 out of 25, and remained at that level for the duration of the study. In other words, participant 27 started off with very good satisfaction with life, and a very low sense of loneliness. It is not surprising, then, that the game experience had little beneficial effect.

For participant 36, total loneliness was again very low throughout, but SWLS was more varied over the study. The participant tells us that the effect of the game on satisfaction with life was minimal, and so the variation is assumed to be due to external factors.
8.8 Summing up

This chapter has described experiment #2, conducted over May-June 2014. The aims of running this experiment were to address issues found in experiment #1, and evaluate whether these corrections resulted in a more robust experimental method and research instrument design. The objectives for this iteration of the experiment were to:

- Provide more of a sense of community to the players,
- Ensure that the players experienced more social interactions within the game,
- More directly measure the emotional impact of playing the game, and
- Find out if cognitive scores would show an improvement if the game was played over a longer period.

The hypotheses tested in this experiment were:

**H1:** Multiplayer online video games contribute to an improvement in cognitive function.

**H2:** Multiplayer online video games contribute to an improvement in subjective loneliness.

**H3:** Multiplayer online video games contribute to an improvement in subjective quality of life.

It is safe to say that none of the hypotheses are conclusively supported by the results. The cognitive test results were mixed, with encouraging results for the Stroop colour-word test not reflected in the other results, and contradicted entirely in the case of the Stroop colour tests. H1 is rejected. The loneliness results show some promise, with one participant showing the ‘ideal SSRD’ shape, specifically for social loneliness. On the other hand, this participant is not brain-injured, and the results for loneliness overall are not significant at

---

Table 8.19. Compliance for participants who provided qualitative responses.

<table>
<thead>
<tr>
<th>Participant ID</th>
<th>Phases in which tests were taken</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>1, 2</td>
</tr>
<tr>
<td>27</td>
<td>1, 2, 3, 4</td>
</tr>
<tr>
<td>32</td>
<td>1, 3, 4</td>
</tr>
<tr>
<td>35</td>
<td>1, 2, 4</td>
</tr>
<tr>
<td>36</td>
<td>1, 2, 3, 4</td>
</tr>
</tbody>
</table>
$p < .05$. H2 is thus not supported, but this researcher does hope that more participants in a larger study would show positive results. The Satisfaction With Life results again did not show a significant improvement, and although one brain-injured participant did show the ‘ideal SSRD’ pattern, H3 is also unsupported.

Qualitative responses were gathered from five participants, and painted a mainly positive picture about how well the game experience created a sense of belonging to a community. But more needs to be done to address the issue of there being no one else online with whom to interact. Indeed, this problem has afflicted commercial games and can be a deciding factor in their success – Steinkuehler and Williams (2006) present the example of Asheron’s Call 2 (Turbine, Inc., 2002-2005), a ‘lonely game’ due to the low number of subscribers. In the present study, daily emails helped to keep players engaged with the game, but participants who scored highly for loneliness at the start tended to take the tests more sporadically. These players arguably have the most to gain from the proposed benefits of multi-player online games. In order to effectively test this, these participants have to be encouraged to take the tests more often, and consequently efforts to improve engagement, such as sending the daily emails, should be made from the start of the study.

In experiment #2, the game design was modified with the intention that players would experience more social interactions, as the number of players logged in at the same time would probably be low. This was achieved by adding a turn-based trading mechanic to the game. Take-up of the new trading mechanic was lower than had been hoped for. It may be that some players are content to be part of the virtual community, without feeling that they must interact with other players – Ducheneaut et al. (2006) note that this is a valid style of enjoying an online multi-player game. On the other hand, players may be hesitant to initiate interactions with people they do not know. If this is the case, future games could provide new ways of interacting that reduce this barrier. The researcher made more trade offers than any of the other players, in an attempt to ‘kick start’ the trading process. This was not a great success, and the conclusion is that the trading mechanic alone is not sufficient to generate a self-sustaining flow of interactions between players. The root of the problem with the mechanic implemented may be that it does not engender a ‘virtuous circle’ or ‘chain reaction’ of interactions between players, and solving this game design issue would be a good next step in boosting social interactions.

Another objective was to more directly measure the emotional impact of playing the game. This was achieved by measuring loneliness and satisfaction with life directly. Experiment #2 was twice the duration of experiment #1, as more time spent in the game could potentially result in greater change to the cognitive test scores. It turned out that the
cognitive test scores did not show appreciable improvement in the second experiment. So, further study is still required to find evidence for cognitive improvement that can be ascribed to playing the multi-player online game.

Overall, the goal of the experiment was to evaluate the experimental method, and game software design for its suitability in a larger, better-resourced study. To conclude, we can say:

- Daily emails were found to be successful in encouraging players to continue to play the game. Sending them from the start of the study could incentivise more participants to comply more fully with the test schedule.
- Further A/B testing could confirm whether or not procedurally generated game events are helpful in encouraging greater participation.
- The study duration was doubled, to test whether more time playing the game would be required for changes in cognitive test scores to be seen. In the event, the Stroop tests did show more variation than in experiment #1, but overall no definitive conclusions can be drawn.
- It is not known if the loneliness and satisfaction with life scores would change more (or less) if the duration was changed: study duration is still a variable for further investigation.
- Participants did not engage with the trading mechanic as much as anticipated. A problem identified in retrospect is that a trading interaction is self-contained, in that it will not lead to further interactions. Future game design work should concentrate on mechanics that cause more interactions.
- Follow-up testing at six months suffered from a lack of participation: only four of the sixteen participants replied. Sample size calculations for a larger study may need to take this into account. Furthermore, these results may be skewed as the participants self-selected to respond.
9 Conclusion

This thesis has proposed that playing online multi-player video games can provide a form of cognitive therapy for brain-injured people.

Millions of people globally live with the effects of acquired brain injury (ABI), with stroke and traumatic injury the leading causes. Isolation is common following such injury. Physical disability can literally isolate; speech and language disabilities do so more subtly. An apparently non-disabled survivor can exhibit personality changes that lead to their marginalisation and exclusion. This research has aimed to find a new form of therapy that has the potential to improve the lives of ABI survivors.

This work has extended the reach of empirical research into the beneficial effects of video games for brain-injured people. The experimental approach of Malec et al. (1984) was taken as a starting point, and extended in three ways. Firstly, instead of a single-player game, a prototype multi-player game was used as the intervention. Secondly, the cognitive measures were extended to include tests of loneliness and satisfaction with life. Thirdly, the measures were integrated into the game software, in principle making it possible to participate from anywhere.

Through an iterative process, a prototype multi-player online game was developed, with the design goal of fostering cooperation, collaboration, and altruistic behaviour in the players. This software, with integrated tests, was used as the research instrument in pilot studies, with both brain-injured and non-brain-injured participants.

The contribution to knowledge has been to propose a new form of therapy, to refine an experimental method, and design a research instrument, evaluating these in case studies.

Chapter 1 listed the objectives of the research. They were:

- From the literature, show a legitimate basis for the proposed therapeutic use of multi-player online video games;
- Establish a research approach for finding evidence for the proposed benefits;
- Develop game and research software as required;
- Conduct experiments and collect data; and
- Disseminate the results in order to stimulate further research in this area.

This chapter evaluates the work against these objectives. The limitations of and implications for the work are discussed, and future work is proposed.
9.1 Conceptual model

The first objective of the research was to develop a conceptual framework for the proposed beneficial effects of multi-player online games. From the rehabilitation literature, we find that isolation is common following brain injury. This could be due to mobility, speech and language deficits, personality changes, and cognitive disability. Weiss (1973) describes loneliness as ‘uniformly distressing’, ‘a chronic distress without redeeming features’. In addition to its unpleasantness, loneliness is linked to cognitive decline and physical ill health (Hawkley and Cacioppo, 2010) – a downward spiral. But a high subjective satisfaction with life is possible for disabled people – the so-called ‘disability paradox’. Jones et al. (2010) found that for brain-injury survivors, reported quality of life was best for those with a strong social network.

Turning to a somewhat different body of literature, there have been many examples of video games being used therapeutically, some dating back to the 1980s. There are numerous cases of therapists using games as a way to bridge a communication gap, in examples involving clients with emotional, behavioural and psychiatric disorders, and learning disabilities. Massively multi-player online games (MMOs) do not yet have such a history of therapeutic use, but studies of the players have found that they enjoy helping each other, meeting, and behaving sociably.

Social loneliness is that form of loneliness associated with the absence of an engaging social network, causing feelings of boredom, aimlessness, and marginality (Weiss, 1973). The connection proposed in this thesis is that the activities of MMO players dovetail with the symptoms of social loneliness, and so might provide relief.

According to Sohlberg and Mateer (2001) improvements in cognitive and emotional abilities will proceed together. If this is indeed the case, then there could be a further link, from the emotional therapy which playing an MMO could provide, to improvements in cognitive skills.

9.2 Research approach

Having established a theoretical basis for multi-player video games as a form of therapy for brain-injured people, the next objective was to establish a research approach for investigating the proposed benefits. The overall approach was experimental and quantitative. The experimental design took as its starting point an earlier study into the cognitive rehabilitation potential of a single-player video game (Malec et al., 1984). The reason for choosing that study was its overlap research aims and practical limitations of the present study.
Obtaining ethical approval for the study involved a review of the medical risks and other potential ill effects of video games. It is fair to say that video games have been criticised on many fronts: there is a small, real, risk of a seizure in a susceptible person. There are ongoing debates on the effects of violent video games, video game ‘addiction’, and the sometimes unwholesome attitudes projected by games.

The recruitment strategy followed the recommendations of Blanton et al. (2006), and consisted mainly of personal visits to support groups in east and central London over a period of around ten months. The goal was to recruit ten brain-injured participants, based on the sample size obtained by Malec et al., and noting that Blanton et al. describe recruitment as one of the more challenging aspects of the research process. Ultimately, the number of brain-injured participants was disappointing as only three were recruited. That said, these three people were thoroughly reliable and committed to the research.

A small sample size precludes a randomised controlled trial (RCT) design, but it was felt from the outset that a sufficient sample size for an RCT was an unrealistic goal. Instead, a single-subject research design (SSRD) was chosen. In an SSRD, time series data is collected over the study period while the intervention is introduced, withdrawn, and then re-introduced.

9.3 Game software design and development

In parallel with the recruitment efforts, game software was developed to use as the research instrument. At the start of the study, developing this software was not a given, but the decision to develop bespoke software meant that all aspects of the game were malleable. As play testing was conducted, and feedback gathered from potential participants, the game changed hugely, in ways that would not be possible with an off-the-shelf game. Further, the game design was tailored to specifically test the effects of cooperative, altruistic behaviour.

In the game, players inhabit a shared environment containing food, treasure, and harmful enemies that deplete players’ health. Eating food restores health, but players can only eat food given to them by other players. Thus players are reliant on each other to eat and stay alive. This scenario is intended as a ‘socially significant’ situation as described by Ray (2004, p. 56), designed to heighten the emotional involvement for players. The adversity and uncertainty in the game world was intended to foster cooperation as this has been seen in many contexts (Andras, Lazarus, and Roberts, 2007).

This prototype game was developed with input from brain-injured play test participants, and modified in the light of comments from other brain-injured people, who were
potentially participants in the study. One of the largest changes in the development of the
game was the switch from 3D to 2D. This was felt necessary after it was found that some
people would find it difficult to process the 3D graphics or could feel nauseous.
The game software was downloadable so that participants could run it in their own home.
This design frees participants from the need to be geographically close to the researcher –
Blanton et al. (2006) found that studies have been affected when participants’ travel was
disrupted. Also, this design is ecologically more valid, as this is the same setting in which
games would normally be played.

9.4 Experiments
The first experiment was conducted in August-September 2013, with the focus on the three
brain-injured participants, but including thirteen other non-brain-injured supporting
players. The participants took a short suite of cognitive tests on five consecutive Fridays.
On alternate weeks, they played the prototype multi-player online game. The tests taken
were the Stroop tests, trail-making tests and a reaction time test, chosen for their broad
sensitivity to cognitive change.
The cognitive skills of the participants as measured by these tests were unaffected by
exposure to playing the online multi-player video game. These results agreed with those of
Malec et al. (1984), who had earlier found that a single-player game appeared not to help
participants score higher on mostly the same tests. A second, more important, finding was
the discovery of a methodological obstacle: participants did not log in at the same time, so
missing out on the social experience central to the proposed benefits.
The second experiment, conducted in 2014, refined the experimental protocol to address
the issues found in the first experiment. The game was redesigned to allow for social
interaction when the players were not online at the same time. Two new measures were
added to the tests that the participants took each week. These were the De Jong Gierveld
Loneliness Scale, and the Satisfaction With Life Scale. The loneliness measure chosen
distinguishes between the two types of loneliness, social and emotional, which were
identified by Weiss (1973).
For experiment #2, the study duration was increased to eight weeks to explore the
possibility that more time would be required for benefits to become apparent. The
cognitive test results showed more variation than in experiment #1, but were again
inconclusive. Also for the loneliness and Satisfaction With Life tests, the hypotheses were
not accepted, but the researcher found some reasons to be hopeful for better results in a
future study.
The sample sizes in these experiments were small, and the results have low external validity. They do point to areas where further research with greater numbers of participants could be focused.

### 9.5 Dissemination

The final objective was to disseminate the work, in order to stimulate research in this area. My hope is that such further research will produce positive results, leading to mainstream acceptance of the core proposal. Parts of this work have been published and presented. I presented the experimental work of 2013 at the Interactive Technologies and Games (ITAG) conference held in Nottingham in 2013, with the paper published in the Journal of Assistive Technologies (Colman, Briggs, Turner, and Good, 2014).

I regard accessibility as an important feature of games, and have published and presented on the following accessibility topics: a paper on ‘one-button’ brain-computer interfaces, presented at DSAI 2010 in Oxford (Colman and Gnanayutham, 2010a), a paper on accessible button interfaces in the International Journal of Web-Based Learning and Teaching Technologies (Colman and Gnanayutham, 2013a), and in a chapter on assistive technologies for brain-injured gamers (Colman and Gnanayutham, 2013b) in the book *Assistive Technologies and Computer Access for Motor Disabilities* (Kouroupetroglou, 2013).

I presented an ethical review of the use of video games as therapy, also at DSAI 2010 (Colman and Gnanayutham, 2010b), the aim being to provide a useful review to other researchers seeking ethical approval for research in this area.

In addition to the above publications, I have presented at University of Portsmouth internal events and elsewhere, and exhibited at conferences. Finally, in 2013 I was invited to participate in the ACM CHI 2013 Doctoral Consortium, held in Paris, and later that year, at the BCS Doctoral Consortium at Brunel University, London.

### 9.6 Contribution to knowledge

This work contributes significantly to the serious games health research community. In the experimental work described, no claims are made for the external validity of the findings, as the sample sizes are far too small. They are to be considered as case studies. Rather, the contribution has been

- Developing a conceptual basis for a new form of therapy,
- Refining an experimental method,
• Design of a research instrument in the form of game software with integrated tests, and
• Evaluation of the method and instrument in case studies.

This work has identified a research area at the intersection of loneliness - specifically the isolation often experienced after brain injury - and the therapeutic potential of online multi-player video games. Research into massively multiplayer online games (MMOs) thus far has laid important groundwork by establishing that players enjoy meeting and helping one another. The current work extends MMO research, proposing them as a potential form of rehabilitation for a neglected group of people who would benefit from exactly the kind of social interaction MMOs provide.

This core proposal needed to be tested experimentally, and this testing has presented challenges. Every brain injury is different, meaning that no two people would experience exactly the same disabilities or personality changes resulting from an ABI. Using a single-subject research design (SSRD) as the experimental method has provided a way forward, even though the number of participants was small. The research instrument - the game software - was made downloadable, and had the experimental measures integrated into it. That removed location and mobility as barriers to participation. Running the experiments described in this thesis have allowed the evaluation and iterative refinement of the software and method, and have paved the way for the future stages of this research.

9.7 Limitations
The experiments described have been pilot studies with very low numbers of participants. The aim of running the experiments has been to evolve an experimental method, choice of measures, and game software design that can be employed in a larger study. In this, I believe the research has been a success. The clear limitation of the experiments is that they do not have the external validity of a study with more participants, and the present results can only be indicators of potential findings in a larger study. This was not an unexpected failing, but a recognised limitation of the study. Other issues were encountered over the course of the study and could only have been discovered by conducting the experiments. These learning experiences should inform the future stages of this research.

The second iteration of the experiment attempted to address the root of the problem discovered in experiment #1. Participants did not have to log in at the same time as social interactions took the form of messages rather than real-time interaction. Under this new design, players could communicate interactively if they were in the game world at the same time, or leave messages and trade in-game commodities if they were not.
Unfortunately, these features were largely not taken up by the participants. On reflection, the best kind of game mechanic to employ here would be one that caused a cascading ‘chain reaction’ of interactions, which the trading game mechanic did not. The upshot is that many players are needed to populate the game world, as only some players will initiate interactions. Dropout and non-compliance rates should also be factored into sample size calculations for future work. In experiment #2, of the sixteen participants, only seven took tests in phase three, and only one participant took the tests at all nine scheduled points.

9.8 Potential impact of the research

As the Internet spreads to ever-remoter parts of the world, a therapy based on it has the potential to reach people who are isolated due to disability or geographical location, where traditional rehabilitation services can be scarce. Hyder et al. (2007) report the increasing prevalence of TBI in low- and middle-income countries, where this is a very real problem. This researcher hopes that social games will be accepted as another game genre that can provide real therapeutic benefits for the players. If it can be shown that multi-player online video games can improve perceived loneliness and subjective satisfaction with life, then many people would have access to a new form of therapy.

What is the likely take-up of video game therapy? It may be that primary healthcare providers would take a conservative approach to any new type of intervention. What is much more likely is that people with an ABI decide for themselves to play such games, as social gaming becomes the norm for more and more people. Popular MMOs have millions of subscribers, and the odds are that a young person who sustains a traumatic brain injury today will already have played online multi-player games. The implication is that there is no need to persuade healthcare agencies to provide online games as a form of therapy. The need is to educate people with brain injury, and their carers and family members, about the potential benefits and risks of these games. It may be that some game genres are more beneficial than others. Such games may place more emphasis on collaboration and altruism, but it could well be that competitive social games make just as much contribution to a strong social network, and so could have the same beneficial effects.

9.9 Future work

The next step is to run a larger study, with a sample size that will enable a randomised controlled trial. The objective of the larger study would be to definitively answer the following questions, as proposed by the conceptual model but not shown by the present findings.
• Can an online multi-player game improve loneliness and subjective wellbeing, and by how much?
• How many people benefit? Of those that do, is there a common characteristic?
• Is social loneliness affected more than emotional loneliness?
• Is loneliness affected more than overall satisfaction with life? If so, are there other facets of satisfaction with life that could perhaps be improved?
• Does the game need to be of a specific genre, or have particular gameplay features?

The RCT would employ a between-subjects design, with three matched groups. One group would play a single-player game, while the second group plays a multi-player game. The control group would not play any video game. This design would test the hypothesis that playing an online multi-player game would result in better test scores than no game at all, or a single-player equivalent game. Having an open source or bespoke game, such as the one developed in the present study, would allow for the single- and multi-player variants to be as close as possible to each other.

Blanton et al. (2006) report that many studies have faltered due to recruitment difficulties. In the present study, recruitment was indeed a challenge. What is the required sample size for the proposed RCT? Sample size calculation is a delicate balancing act: many studies make type II errors due to a sample size that is too small (Moher, Dulberg, and Wells, 1994), but having too many participants would waste their time and effort, and potentially expose them to risk unnecessarily. Section 3.1.2 details methods for calculating and estimating sample size. A precise figure can be calculated in principle, but only if precise values are known for the formula variables; alternatively, graphical devices such as Altman’s (1980) nomogram might be used.

The dropout and compliance rates should also be taken into account. It is possible that the compliance rate could be improved by modifying the game software. For instance, the loneliness and satisfaction with life questionnaires, and any other future measures, could be ‘gamified’, perhaps by having in-game characters ask the questions. Players could be incentivised to participate by being rewarded with in-game items or commodities. This idea was not implemented in the present study, and could have made a difference to the compliance level.

A further consideration, specific to this research, is to estimate the number of players we need to adequately populate the game world. A certain minimum number of players should be online at any given moment, so that all players experience social interactions. The login data from the present study can be extrapolated to give a rough estimate of the numbers
required. In phase 2 of experiment #2 (the first ‘game’ phase), the 16 participants logged a total of 24 hours online. The phase length in experiment #2 was two weeks, giving 0.1 hours of login time per participant per day. If we wanted, say, ten players online at all times over a ten-hour period of each day, we would need 1000 players! Finding ways to engage the players, and so increasing mean login time, is a crucial step in reducing the number of players required. A/B tests with small groups of participants could be used to compare incentives and game mechanics, to find those that are most effective. The trading game mechanic explored in experiment #2 was designed to increase interactions between players, reducing the need to have so many players online at the same time. Further development of this kind of ‘offline’, turn-based feature could bear fruit. The best kind of mechanic would generate multiple interactions, each of which could lead to further interactions, and so on, in a ‘chain reaction’.

Recruiting a much larger group of participants than has been achieved to date requires a different strategy. A step change in the number of players could potentially be achieved by employing more widespread advertising than was used in the present study. Offering the game client software on mobile platforms as well as Windows PC would increase the pool of potential participants.

Let us assume that an RCT such as that described is conducted, and does indeed show that a multi-player online game does reduce loneliness and improve satisfaction with life. Further work should then seek to understand the essential features of multi-player online games that provide beneficial effects. As before, open source software would enable the same game to have multiple variants developed, so individual game features could be evaluated. The game software could be set up to support A/B testing, where half the players get a new feature or setting, while the other half do not. This will allow for comparisons between game features, but would only work up to a point: it would not be fair to have half the players behaving altruistically while the other half fight them competitively. This being the case, multiple experiments using different game types would be needed.

9.10 Concluding remarks

Sadly, traumatic brain injury and stroke are not going away. The next two decades are predicted to see a tripling in stroke mortality in low- and middle-income countries (World Heart Federation, 2016), and the incidence of stroke in the US could increase 20% by 2030 (Ovbiagele et al., 2013). This thesis has proposed a form of therapy for brain injury
survivors, complementary to other rehabilitation approaches, which could be deployed anywhere the Internet reaches.

Green and Bevelier (2006) have noted the challenges of this kind of research area, and that it is not easy to produce perceptual or cognitive changes through video games. Indeed, whether or not specifically designed ‘brain training’ software has any real effect is an area of on-going research and debate, with an unenthusiastic position statement (Max Planck Institute for Human Development and Stanford Center on Longevity, 2014) hotly contested by the industry.

The aim of this research has been to help people with acquired brain injury, by investigating the potential of multi-player online video games as a form of therapy. There is still much more work to do. Every person who survives a brain injury will be affected in a different way, and respond differently to any given therapy. The benefits of online multi-player video games proposed in this thesis have yet to be proven, but the present work paves the way forward for doing so. My hope is that the current work can guide future studies, leading to mainstream acceptance of the proposed benefits of multi-player online video games, and, ultimately, an improvement in quality of life for brain-injured people.
References


Namco (1979) Galaxian [Video game]. Chicago, IL: Midway Games.


Taito Corporation (1978) *Space Invaders* [Video game]. Chicago, IL: Midway Games.


Appendix 1  Information sheet and consent form for experiment #1

Multi-player online video games for cognitive rehabilitation
Researcher: Jason Colman, student at University of Portsmouth
Supervisors: Dr Jim Briggs, Dr Louise Turner, Dr Alice Good

My email:  jason.colman@port.ac.uk
Web page:  www.amju.com/mygame
My Phone:  07902 454 279

My name is Jason Colman and I am a student at the University of Portsmouth. I am conducting a study to find out if playing an online multi-player video game could help people with a brain injury.

The potential benefit of the research is that it may help you to improve your skills in areas such as communication, memory, and planning. Overall, multi-player online games may be identified as a form of therapy, which could be used to help other people who have suffered a brain injury.

Who can take part?

I am looking for people who are over 18 years old, who have had a stroke or traumatic brain injury, who are interested in playing video games. You should not already be regularly playing a multi-player online video game.

You need a computer (Windows PC or Mac) which is connected to the internet. You need to be able to read the screen and be able to move and click the mouse cursor, using any controller that suits you.

You need to be able to give your consent to participating (there is a consent form at the end of this document and on the web site).

What you will be asked to do

You will be asked to download a special video game and play for a few hours, spread over 4 weeks. You will be asked to play on specific days.

In the game, you can chat to other players. Only invited players are able to log in to the game, so you would only be playing with other people who have kindly agreed to help with this research project.

There are tests in the game, which you are asked to take, which will be used to measure the beneficial effect of playing the game.
Potential risks or negative effects associated with participation

Staring intently at a screen can cause visual fatigue, and flashing images can cause seizures in people susceptible to photosensitive epilepsy. (This is a rare condition.) To minimise these risks, gaming sessions will be kept short and regular breaks scheduled.

Video games have been criticised for their violent content. Some games contain other offensive content. Some researchers consider video game ‘addiction’ to be a risk for gamers. The game you will be playing will not have violent content.

What if I want to withdraw from the study?

You may at any time withdraw from the study. You do not have to give any reason, and no one can attempt to dissuade you.

Confidentiality and privacy

Your identity will be kept confidential. Your written permission would be sought before publishing information which could identify you. Material which could not be used to identify you may be published or presented with the aim of benefiting others.

Your email address is used as a way of identifying you when you first log in to the game – but it will not be shared with anyone else.

Dates

Currently the plan is for the study to run during the period July-September 2013.

How can I join in?

If you would like to participate, Please send me an email to let me know of your interest. My email address is jason.colman@port.ac.uk.

Before you can participate, you must give your consent. Please can you print off and sign the consent form below. (If you are unable to sign, I am allowed to accept verbal consent in front of a witness.)

Once I have your consent, I will email you a link to download the special video game, with instructions.

Any questions?

If you have any questions about this research project, now or in the future, please feel free to contact me or my first supervisor:

Jason Colman
Jason.Colman@port.ac.uk
Phone: 07902 454279

Dr Jim Briggs
Jim.Briggs@port.ac.uk
CONSENT FORM

Multi-player online video games for cognitive rehabilitation

Researcher: Jason Colman
jason.colman@port.ac.uk
Phone: 07902 454279

Please initial each numbered item if content

1. I confirm that I have read and understood the attached information sheet for the above study. I confirm that I have had the opportunity to consider the information, ask questions and that these have been answered satisfactorily.
2. I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason.
3. I agree to take part in this study.

Name of participant:

Your email address (please print clearly):

Mobile phone number (so I can send you reminder texts, if that is OK with you)

Please indicate the most convenient time(s) for you to play this game:
(E.g. morning, afternoon, evening, or any times you definitely can’t do it)

Today’s date:

Signature:

Now please send this form to:

Jason Colman (PhD Student)
School of Computing
Buckingham Building
Lion Terrace, Portsmouth, PO1 3HE

...or you can scan it and email it to me, if you prefer.

Thank you for helping with this research project!
Appendix 2  Information sheet and consent form for experiment #2

My name is Jason Colman and I am a student at the University of Portsmouth. I am conducting a study to find out if playing an online multi-player video game could help people with a brain injury.

The potential benefit of the research is that it may help you to improve your skills in areas such as communication, memory, and planning. Overall, multi-player online games may be identified as a form of therapy, which could be used to help other people who have suffered a brain injury.

Who can take part?

I am looking for people who are over 18 years old, who have had a stroke or traumatic brain injury, who are interested in playing video games. You should not already be regularly playing a multi-player online video game.

You need a computer (Windows PC) which is connected to the internet. You need to be able to read the screen and be able to move and click the mouse cursor, using any controller that suits you.

You need to be able to give your consent to participating (there is a consent form at the end of this document and on the web site).

What you will be asked to do

You will be asked to download a special video game and play for a few hours, spread over 8 weeks from the beginning of May 2014. You will only be asked to play on some days.

In the game, you can chat to other players. Only invited players are able to log in to the game, so you would only be playing with other people who have also kindly agreed to help with this research project.

There are tests in the game, which you are asked to take, which will be used to measure the beneficial effect of playing the game. You are also asked questions about how you feel. Please do answer these questions openly and honestly as far as you can. All your answers are confidential.
Potential risks or negative effects associated with participation

Staring intently at a screen can cause visual fatigue, and flashing images can cause seizures in people susceptible to photosensitive epilepsy. (This is a rare condition.) To minimise these risks, gaming sessions will be kept short and regular breaks scheduled.

Video games have been criticised for their violent content. Some games contain other offensive content. Some researchers consider video game ‘addiction’ to be a risk for gamers. The game you will be playing will not have violent content.

What if I want to withdraw from the study?

You may at any time withdraw from the study. You do not have to give any reason, and no one can attempt to dissuade you. Please note that I would like your permission to use your results even if you withdraw. If this is not OK, please just let me know.

Confidentiality and privacy

Your identity will be kept confidential. Information which could identify you will not be published. Material which could not be used to identify you may be published or presented with the aim of benefiting others.

Your email address is used as a way of identifying you when you first log in to the game – but it will not be shared with anyone else.

Dates

The study will run during May and June 2014.

How can I join in?

If you would like to participate, please send me an email to let me know of your interest. My email address is Jason.Colman@port.ac.uk.

Before you can participate, you must give your consent. Please can you print off and sign the consent form below. (If you are unable to sign, I am allowed to accept verbal consent in front of a witness.)

Once I have your consent, I will email you a link to download the special video game, with instructions.

Any questions?

If you have any questions about this research project, now or in the future, please feel free to contact me or my first supervisor:

Jason Colman
Jason.Colman@port.ac.uk
Phone: 07902 454279

Dr Jim Briggs
Jim.Briggs@port.ac.uk
CONSENT FORM

Multi-player online video games for cognitive rehabilitation

Researcher: Jason Colman
jason.colman@port.ac.uk
Phone: 07902 454279

Please make sure you agree with these points:

1. I confirm that I have read and understood the attached information sheet for the above study. I confirm that I have had the opportunity to consider the information, ask questions and that these have been answered satisfactorily.
2. I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason.
3. If I withdraw from the study, my results would still be retained and used, unless I ask for the results to be destroyed.
4. I agree to take part in this study.

Name of participant:

Your email address (please print clearly):

Today's date:

Signature:

Now please send this form to:

Jason Colman (PhD Student)
School of Computing
Buckingham Building
Lion Terrace, Portsmouth, PO1 3HE

..or you can scan it and email it to me, if you prefer.

Thank you for helping with this research project!
Appendix 3  In-game trades made in experiment #2

The table in this appendix shows all the trades initiated in the prototype multi-player online game in experiment #2, detailed in chapter 8.

Table A3.1. In-game trades initiated during experiment #2.

<table>
<thead>
<tr>
<th>Date trade offered</th>
<th>Trade offered by (Player ID)</th>
<th>Trade offered to (Player ID)</th>
<th>Initiated by researcher?</th>
<th>Accepted?</th>
</tr>
</thead>
<tbody>
<tr>
<td>23/05/2014</td>
<td>35</td>
<td>7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>24/05/2014</td>
<td>12</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>27/05/2014</td>
<td>1</td>
<td>13</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>27/05/2014</td>
<td>1</td>
<td>15</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>27/05/2014</td>
<td>1</td>
<td>30</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>27/05/2014</td>
<td>1</td>
<td>32</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>27/05/2014</td>
<td>1</td>
<td>33</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>27/05/2014</td>
<td>1</td>
<td>34</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>27/05/2014</td>
<td>1</td>
<td>35</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>27/05/2014</td>
<td>1</td>
<td>36</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>27/05/2014</td>
<td>1</td>
<td>41</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>27/05/2014</td>
<td>1</td>
<td>12</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>27/05/2014</td>
<td>24</td>
<td>15</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>27/05/2014</td>
<td>32</td>
<td>1</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>29/05/2014</td>
<td>35</td>
<td>1</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>30/05/2014</td>
<td>2</td>
<td>35</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>30/05/2014</td>
<td>2</td>
<td>33</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>30/05/2014</td>
<td>2</td>
<td>32</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>30/05/2014</td>
<td>2</td>
<td>29</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>30/05/2014</td>
<td>2</td>
<td>28</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>30/05/2014</td>
<td>2</td>
<td>27</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>30/05/2014</td>
<td>2</td>
<td>26</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>30/05/2014</td>
<td>2</td>
<td>25</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>30/05/2014</td>
<td>2</td>
<td>24</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>30/05/2014</td>
<td>2</td>
<td>23</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>30/05/2014</td>
<td>2</td>
<td>22</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>30/05/2014</td>
<td>2</td>
<td>15</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>30/05/2014</td>
<td>2</td>
<td>13</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>30/05/2014</td>
<td>2</td>
<td>12</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>30/05/2014</td>
<td>2</td>
<td>11</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>30/05/2014</td>
<td>2</td>
<td>50</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>01/06/2014</td>
<td>35</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>02/06/2014</td>
<td>15</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>05/06/2014</td>
<td>34</td>
<td>1</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>05/06/2014</td>
<td>34</td>
<td>12</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix 4  Ethical Approval and Form UPR16 (Research Ethics Review Checklist)

Ethical approval was confirmed by emails from Dr John Williams, Chair of the Faculty of Technology Ethics Committee at the University of Portsmouth, dated 14 March 2012 and 23 May 2014.
FORM UPR16
Research Ethics Review Checklist

Please include this completed form as an appendix to your thesis (see the Postgraduate Research Student Handbook for more information)

<table>
<thead>
<tr>
<th>Postgraduate Research Student (PGRS) Information</th>
<th>Student ID:</th>
<th>502355</th>
</tr>
</thead>
<tbody>
<tr>
<td>PGRS Name: Jason Colman</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Department: School of Computing</td>
<td></td>
<td></td>
</tr>
<tr>
<td>First Supervisor: Dr Jim Briggs</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Start Date: 1 Feb 2010</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Mode and Route: Part-time</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Title of Thesis: Multi-player online video games for cognitive rehabilitation for the brain injured</td>
<td>45 670</td>
<td></td>
</tr>
<tr>
<td>Thesis Word Count: (excluding ancillary data)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

If you are unsure about any of the following, please contact the local representative on your Faculty Ethics Committee for advice. Please note that it is your responsibility to follow the University’s Ethics Policy and any relevant University, academic or professional guidelines in the conduct of your study. Although the Ethics Committee may have given your study a favourable opinion, the final responsibility for the ethical conduct of this work lies with the researcher(s).

UKRIO Finished Research Checklist:
(If you would like to know more about the checklist, please see your Faculty or Departmental Ethics Committee rep or see the online version of the full checklist at: http://www.ukrio.org/what-we-do/code-of-practice-for-research/)

a) Have all of your research and findings been reported accurately, honestly and within a reasonable time frame? YES ☒ NO ☐

b) Have all contributions to knowledge been acknowledged? YES ☒ NO ☐

c) Have you complied with all agreements relating to intellectual property, publication and authorship? YES ☒ NO ☐

d) Has your research data been retained in a secure and accessible form and will it remain so for the required duration? YES ☒ NO ☐

e) Does your research comply with all legal, ethical, and contractual requirements? YES ☒ NO ☐

Candidate Statement:
I have considered the ethical dimensions of the above named research project, and have successfully obtained the necessary ethical approval(s)

Ethical review number(s) from Faculty Ethics Committee (or from NRES/SCREC): See below

If you have not submitted your work for ethical review, and/or you have answered ‘No’ to one or more of questions a) to e), please explain below why this is so:

Approval emails from Dr John Williams, 14 March 2012 and 23 May 2014

Signed (PGRS): Date:

UPR16 – August 2015