Games as personality profiling tools

Conference Paper · August 2011

DOI: 10.1109/CIG.2011.6032007

Giel van Lankveld
Tilburg University

H. Jaap Van Den Herik
Leiden University

Arnoud Arntz
University of Amsterdam

Quantifying Individual Player Differences

Quantifying Individual Player Differences

PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan Tilburg University,
op gezag van de rector magnificus,
prof. dr. Ph. Eijlander,
in het openbaar te verdedigen ten overstaan van een
door het college voor promoties aangewezen commissie
in de aula van de Universiteit
op woensdag 27 februari 2013 om 16:15 uur

door

Giel van Lankveld
geboren op 3 april 1982 te Gemert

Promotores:
Prof. dr. H.J. van den Herik
Prof. dr. A.R. Arntz

Copromotor:
Dr. ir. P.H.M. Spronck

Leden van de beoordelingscommissie:
Prof. dr. M.H.J. Bekker
Prof. dr. A. Plaat
Prof. dr. E.O. Postma
Prof. dr. V. Evers
Prof. dr. H. Iida
Prof. dr. S.M. Lucas

This research has been partly funded by the project KIM (Kennis in Modellen) of the Korps Landelijke Politie Diensten (KLPD), supported by the Dutch Ministry of Security and Justice.

SIKS Dissertation Series No. 2013-[NUMMER]
The research reported in this thesis has been carried out under the auspices of SIKS, the Dutch Research School for Information and Knowledge Systems.

TiCC Ph.D. Series No. 25

ISBN ?-?-?-?-?

Preface

Modern video games have enjoyed over 60 years of progress. From the first simple game in 1947[1] to the complex commercial games of today, the medium of video games has come a long way. Today, we can play games producing screenshots which are visually indistinguishable from photographs. Games feature literally thousands of characters at once; entire cities like ancient Rome and modern-day New York have been carefully reproduced in order to provide playgrounds for gamers. The future of video games promises even greater developments.

Research in video games has been moving forward steadily alongside developments in the games industry. Indeed, we have seen that artificial intelligence research in games started in the area of classical games like chess, but nowadays games research is quite broad and focusses on areas such as the movement of game characters, selecting adequate difficulty for the player, modelling players' mental traits, and searching through game decision trees.

Our contribution to the field of game research is in improving player modelling with a scientific eye on incongruity theory and personality theory. We have pushed psychology research forward by using games as personality assessment tools.

I am grateful to my supervisor Jaap van den Herik for his guidance during my Ph.D. research and for his detailed advice during the writing of this thesis, and to my second supervisor Arnoud Arntz for his advice on the psychological aspects of my thesis. My sincere thanks goes out to my daily advisor Pieter Spronck for his guidance during research, writing, and his advice during many challenges in my Ph.D. career. Thanks also go to Carel van Wijk for his methodological and statistical advice. Moreover, I would like to thank the students with whom I have collaborated on the research in this thesis: Sonny Schreurs, Iris Balemans, and Evi Joosten. Here, I would like to recognise the effort of the members of the assessment committee for reading and assessing my thesis.

During my Ph.D. research I have had the pleasure to work at two universities. Both Maastricht University and Tilburg University have provided me with a working environment in which I could grow as a researcher. I have also had the great pleasure to work several months as visiting researcher at JAIST[2] in Japan. The experiences have been amazing and I am very thankful to Professor Hiroyuki Iida to have had these opportunities.

My Ph.D. trajectory has, without any doubt, been the most challenging phase of my life so far and I could not have succeeded without the support of those around me. Most importantly, my partner Madelon Maijers has supported me through the ups and downs,

[1]The Cathode ray tube Amusement Device by Goldsmith

[2]Japan Advanced Institute for Science and Technology

for which I am deeply grateful. Furthermore, I would like to thank my colleagues who were there with good advice, a critical eye, and of course a nice cup of coffee along the way.

I would also like to thank my friends who had to suffer from listening to all my rants about research and thesis writing. You all know how much I have enjoyed our conversations, intellectually and otherwise.

Finally, I would like to thank my family: my dad who has given me much good advice and support, my mom who was always caring about me, and of course my brother for our talks about games and life in general.

Giel van Lankveld,
Maastricht, January 2, 2013

Contents

Preface i

Contents iii

List of abbreviations ix

List of definitions xi

List of figures xiii

List of tables xv

1 Introduction 1
1.1 Human players and computer players . . . 1
1.2 Problem statement and research questions . . . 3
1.3 Methodology . . . 4
1.3.1 Questionnaires . . . 5
1.3.2 Techniques . . . 6
1.3.3 Statistical techniques . . . 7
1.3.4 Overview of methodologies used . . . 7
1.4 Thesis structure . . . 8

2 Theoretical background 11
2.1 Modelling users . . . 11
2.2 Modelling opponents . . . 12
2.3 The entertainment value of games . . . 13
2.3.1 Emotional experience in games . . . 14
2.3.2 Entertainment as subjective experience . . . 14
2.3.3 Entertainment and the experience of positive emotion . . . 16
2.3.4 The role of attitudes in entertainment . . . 18
2.3.5 Alternative reasons to play games . . . 18
2.3.6 Methods to evaluate entertainment . . . 19
2.4 Modelling players . . . 20
2.4.1 Modelling player actions . . . 21

2.4.2 Skill-related player differences . . . 22
2.4.3 Player types . . . 23
2.5 Psychological modelling . . . 23
2.5.1 Personality theory . . . 24
2.5.2 The origin of personality theory . . . 25
2.5.3 The five factor model . . . 25
2.5.4 The NEO-PI-R . . . 25
2.5.5 Contemporary methods for measuring personality . . . 27
2.6 Chapter summary . . . 28

3 Incongruity in games **29**
3.1 Incongruity . . . 29
3.1.1 Incongruity theory . . . 30
3.1.2 Incongruity and emotions . . . 31
3.1.3 Origins of incongruity theory . . . 32
3.1.4 Modern uses of incongruity theory . . . 32
3.2 The game: glove . . . 33
3.2.1 Game world . . . 33
3.2.2 The knight . . . 34
3.2.3 The enemies . . . 35
3.2.4 Difficulty and incongruity . . . 35
3.3 Experimental setup . . . 36
3.3.1 Participants . . . 36
3.3.2 Procedure . . . 36
3.3.3 Questionnaire . . . 36
3.3.4 Statistical techniques . . . 37
3.4 Results . . . 37
3.5 Discussion . . . 38
3.6 Chapter conclusions . . . 39

4 Extraversion in games **41**
4.1 Extraversion . . . 41
4.1.1 Game observations . . . 42
4.1.2 Player modelling versus player profiling . . . 43
4.2 The game: letter to the king . . . 43
4.2.1 Story . . . 43
4.2.2 Controls . . . 43
4.2.3 In-game elements . . . 44
4.3 Experimental setup . . . 48
4.3.1 The extraversion experiment . . . 49
4.3.2 Participants . . . 49
4.3.3 Procedure . . . 49
4.3.4 Statistical techniques . . . 49
4.4 Results . . . 49
4.5 Discussion . . . 52

4.5.1 Extraversion . . . 52
4.5.2 Significance level and control variables . . . 52
4.6 Chapter conclusions . . . 53

5 Data-driven personality in games **55**
5.1 Measuring personality . . . 55
5.2 The game: the poisoned lake . . . 56
5.2.1 Controls . . . 56
5.2.2 The game world . . . 56
5.2.3 The story . . . 58
5.3 Experimental setup . . . 58
5.3.1 Participants and time frame . . . 58
5.3.2 Game variables . . . 59
5.4 Results . . . 62
5.5 Discussion . . . 64
5.5.1 Correlations per trait . . . 64
5.5.2 The effects of openness . . . 67
5.5.3 Interpreting the correlations . . . 67
5.5.4 The modelling of extraversion . . . 67
5.5.5 Interpreting regression results . . . 68
5.5.6 Effect sizes and significance . . . 68
5.5.7 Similarities between conversations and multiple-choice questions . . . 68
5.6 Chapter conclusions . . . 69

6 Theory-driven personality in games **71**
6.1 Alternatives to data-driven models . . . 71
6.2 A theory-driven model . . . 72
6.3 Experimental setup . . . 74
6.4 Results . . . 75
6.5 Discussion . . . 76
6.6 Chapter conclusions . . . 77

7 Validating personality in games **79**
7.1 Profiling by using a commercial game . . . 79
7.2 The game: vault 101 . . . 80
7.2.1 General game information . . . 80
7.2.2 Story . . . 80
7.2.3 Setting . . . 81
7.2.4 Controls . . . 81
7.3 Experimental setup . . . 81
7.3.1 Participants and procedure . . . 82
7.3.2 Data-driven variables . . . 83
7.3.3 Theory-driven variables . . . 84
7.3.4 Questionnaires . . . 85
7.3.5 Analysis . . . 86

7.4 Results 86
7.4.1 Data-driven results 86
7.4.2 Theory-driven results 88
7.5 Discussion 88
7.5.1 Comparison of data-driven models 88
7.5.2 Comparison of theory-driven models 89
7.5.3 Explaining the differences in effects 90
7.5.4 General remarks on personality traits 91
7.5.5 Control variables 92
7.6 Chapter conclusions 92

8 Modelling in games **95**
8.1 Player characteristics 95
8.1.1 Personality 96
8.1.2 Intelligence 96
8.1.3 Demographics 96
8.1.4 Skill 97
8.1.5 Preferences 97
8.2 Player responses 97
8.2.1 Behaviour 97
8.2.2 Cognition 98
8.2.3 Emotion 99
8.3 Player environment 99
8.4 Game psychometrics 99
8.4.1 Spectrum of behaviour 100
8.4.2 Low-level actions, high-level implications 100
8.4.3 Experimental control 100
8.4.4 Multiple different game situations 100
8.4.5 Replay value and repeated measurements 101
8.4.6 Comparison to actual behaviour 101
8.4.7 Individual effects versus mass effects 101
8.4.8 Validity 102
8.4.9 Control variables 102
8.5 Application of personality theory in games 102
8.5.1 Games as personality tests 102
8.5.2 Personality-based game adaptation 103
8.6 Pitfalls in applying psychological theory to games 103
8.7 Chapter conclusions 104

9 Conclusions **105**
9.1 Answering the research questions 105
9.2 The problem statement 107
9.3 Future research 107

References **109**

Appendices **118**

A Incongruity questionnaire **119**

B Balancing glove **121**

C Glove difficulty algorithm **125**

D Incongruity contrast analysis **127**

E Extraversion in games instruction booklet **129**

F NWN study descriptives **135**

G NWN study coefficients **143**

H NWN study significant correlations **147**

I Fallout instruction sheet **151**

J Demographics questionnaire **153**

K Fallout study dialog options **155**

L Fallout study descriptives **163**

M Fallout study significant correlations **169**

N Interview validation of game personality **173**
N.1 The game: the poisoned lake (continued) . . . 174
N.2 Experimental setup . . . 174
N.2.1 Participants . . . 174
N.2.2 Procedure . . . 174
N.2.3 Measures . . . 175
N.3 Results . . . 176
N.3.1 Variable computation . . . 177
N.3.2 Descriptives . . . 177
N.3.3 Correlation analysis . . . 178
N.3.4 Combined datasets . . . 178
N.3.5 Reliabilities . . . 180
N.4 Discussion . . . 180
N.5 Appendix conclusions . . . 181

O Behavioural criteria **183**

Summary **185**

Samenvatting **189**

Curriculum vitae **193**

Publications **195**

SIKS dissertation series **197**

TiCC Ph.D. series **209**

List of abbreviations

ANOVA	Analysis of variance
AI	Artificial intelligence
Ass	Assertiveness
AMME	Automatic mental model evaluation
DDA	Dynamic difficulty adjustment
EQ	Emotion questionnaire
Exc	Excitement seeking
ESM	Experience sampling method
FFM	Five factor model
GECK	Garden of eden creation kit
GQ	General information questionnaire
G.O.A.T.	Generalized occupational aptitude test
Gre	Gregariousness
HCI	Human computer interaction
IQ	Incongruity questionnaire
MANOVA	Multiple analysis of variance
NEO-FFI	Neuroticism, extraversion, and openness five factor inventory
NEO-PI	Neuroticism, extraversion, and openness personality inventory
NEO-PI-R	Neuroticism, extraversion, and openness personality inventory revised
NWN	Neverwinter Nights
NC	New criteria
NCA	New criteria agreeableness
NCC	New criteria conscientiousness
NCE	New criteria extraversion
NCN	New criteria neuroticism
NCO	New criteria openness
NPC	Non-player character
OCEAN	Openness conscientiousness extraversion agreeableness neuroticism
PQ	Personality questionnaire
PBL	Player behaviour logging
PM	Player model
Pos	Positive emotion
RVO	Rated video observation

PS	Problem statement
RTS	Real-time strategy
RQ	Research question
TAM	Technology acceptance model
TRA	Theory of reasoned action
UM	User model
War	Warmth

List of definitions

1.1 Definition (Psychological models) 2
1.2 Definition (Player modelling) 2
1.3 Definition (Personality profiling) 2
1.4 Definition (Incongruity) 6
1.5 Definition (Emotion) 6
1.6 Definition (Valence) 6

2.1 Definition (User model) 11
2.2 Definition (Preferences) 12
2.3 Definition (Entertainment) 14
2.4 Definition (Entertaining experience) 15
2.5 Definition (Well shaped difficulty curve) 15
2.6 Definition (Dynamic difficulty adjustment) 16
2.7 Definition (Implicit optimisation) 16
2.8 Definition (Explicit optimisation) 16
2.9 Definition (Intrinsically rewarding) 17
2.10 Definition (Flow) 17
2.11 Definition (Player modelling) 20
2.12 Definition (Player model) 21
2.13 Definition (Action model) 21
2.14 Definition (Opponent model) 21
2.15 Definition (Skill) 23
2.16 Definition (Skillful) 23
2.17 Definition (Player type) 23
2.18 Definition (Personality) 24
2.19 Definition (Personality theory) 24

3.1 Definition (Incongruity theory) 29
3.2 Definition (Boredom) 31
3.3 Definition (Frustration) 31
3.4 Definition (Pleasure) 31
3.5 Definition (Difficulty balancing) 33
3.6 Definition (Easy) 35
3.7 Definition (Balanced) 35

3.8 Definition (Hard) 35

4.1 Definition (Extraversion) 42
4.2 Definition (Player profiling) 43
4.3 Definition (In-game element) 44

5.1 Definition (Automated personality profiling) 56
5.2 Definition (Z-scores) 60
5.3 Definition (Pooled variable) 60
5.4 Definition (Unpooled variable) 60

6.1 Definition (Behavioural criteria) 73
6.2 Definition (Operationalisation) 73

7.1 Definition (Ecological validity) 79

8.1 Definition (Model) 95
8.2 Definition (Framework) 95
8.3 Definition (General behaviour) 98
8.4 Definition (Low-level behaviour) 98
8.5 Definition (High-level behaviour) 98
8.6 Definition (Cognition) 98

List of figures

2.1 The circumplex model of affect. 15
2.2 Schematic representation of flow. 17

3.1 Schematic representation of incongruity. 30
3.2 The game GLOVE. On the left side we see the knight. On the right from top to bottom, we see the dragon, the ninja, and the witch. 34
3.3 Incongruity questionnaire results. 38

5.1 The five outside areas of the game world. 57
5.2 A screenshot of a typical conversation. 61
5.3 The interior of a house containing three movement triggers. 62

7.1 A screenshot of a conversation in FALLOUT 3. 84

List of tables

1.1 Questionnaires and techniques used. 5
1.2 Methodology matrix. 8
1.3 Thesis structure. 9

3.1 Descriptive statistics for each Glove difficulty. 37

4.1 Correlations between NEO-PI-R scores and game items. 50
4.2 Control questions. 51
4.3 Correlations between control questions and game items. 51

5.1 Characters encountered 57
5.2 Descriptive statistics. 62
5.3 Stepwise linear regression between traits and game variables. 63
5.4 Total number of positive (pos) and negative (neg) correlations per group. . . 64
5.5 Highest and the lowest correlations for each group. 65

6.1 Theory-driven model variables. 74
6.2 Descriptive statistics. 75
6.3 Correlations. 76

7.1 Names and descriptions of the NPCs. 81
7.2 Theory-driven model variables. 85
7.3 Descriptive statistics of NEO-FFI scores. 87
7.4 Stepwise linear regression analysis of traits and game variables. 87
7.5 Positive and negative correlations per variable group. 88
7.6 Minimum and maximum correlations per group. 89
7.7 Descriptives of the theory-driven variables. 90
7.8 Correlations between theory-driven variables and personality traits. 90
7.9 Comparison between Chapter 5 and Chapter 7 R^2 scores. 91
7.10 Correlations between the control variables themselves. 93

A.1 The incongruity questionnaire 120

D.1 Contrasts 128

F.1 NWN study game variables: descriptive statistics 135
F.1 NWN study game variables: descriptive statistics 136
F.1 NWN study game variables: descriptive statistics 137
F.1 NWN study game variables: descriptive statistics 138
F.1 NWN study game variables: descriptive statistics 139
F.1 NWN study game variables: descriptive statistics 140
F.1 NWN study game variables: descriptive statistics 141
F.1 NWN study game variables: descriptive statistics 142

G.1 NWN study: regression coefficients openness. 143
G.2 NWN study: regression coefficients conscientiousness. 144
G.3 NWN study: regression coefficients extraversion. 144
G.4 NWN study: regression coefficients agreeableness. 145
G.5 NWN study: regression coefficients neuroticism. 146

H.1 NWN study: significant correlations 147
H.1 NWN study: significant correlations 148
H.1 NWN study: significant correlations 149
H.1 NWN study: significant correlations 150

K.1 Fallout study: conversation dialog 155
K.1 Fallout study: conversation dialog 156
K.1 Fallout study: conversation dialog 157
K.1 Fallout study: conversation dialog 158
K.1 Fallout study: conversation dialog 159
K.1 Fallout study: conversation dialog 160
K.1 Fallout study: conversation dialog 161
K.1 Fallout study: conversation dialog 162

L.1 Fallout study game variables: descriptive statistics 163
L.1 Fallout study game variables: descriptive statistics 164
L.1 Fallout study game variables: descriptive statistics 165
L.1 Fallout study game variables: descriptive statistics 166
L.1 Fallout study game variables: descriptive statistics 167
L.1 Fallout study game variables: descriptive statistics 168

M.1 Fallout study: correlations 169
M.1 Fallout study: correlations 170
M.1 Fallout study: correlations 171

N.1 Video interview protocol. 176
N.2 Back and Egloff (2009) interview descriptives. 177
N.3 NEO-PI-R trait descriptives. 177
N.4 Correlations between the NEO-PI-R trait scores and the interview trait scores.178
N.5 NEO-PI-R - Interview correlations. 178
N.6 Linear regression analysis for game variables on interview FFM scores. 179

N.7 Correlational between game, NEO-PI-R, and interview scores. 179
N.8 Interview item reliabilies found for each trait. 180

O.1 Behavioural criteria selected from the Back and Egloff (2009) experiment. . . 184

Chapter 1

Introduction

Players differ in playing strength. In some games, like chess, we honour good players by awarding titles, such as World Champion, International Grand Master, and International Master. In chess, the qualification of differences in playing strength is based on results. Results are expected to reflect the player's understanding of the game. In contrast to chess, in video games the quantification of the individual differences relies mostly on the behaviour of the players[1]. A player's behaviour is guided by three processes: (1) cognition (e.g., the player's thinking during play), (2) perception (e.g., the player's observations), and (3) the capability with which the player handles the computer and the program.

In video games we observe a player's behaviour by looking at his input (e.g., mouse clicks and keystrokes). In real life, we can inspect a much larger range of behaviour. Behavioural observation in real life has guided most of the discoveries in the field of psychology. In order to be successful at computer observation we need to consider carefully what we are observing. For instance, a player's actions can be understood and become meaningful if we relate them to the context in which they are performed. Using collected data we may be able to construct models that help us determine the characteristics of the player. The main aim of this thesis is to develop methods by which we can accurately and automatically quantify individual player differences.

In Section 1.1 we introduce the modelling of human and computer players. In Section 1.2 we formulate a problem statement and the corresponding research questions. In Section 1.3 we give an overview of the research methods used and the chapters they apply to. Finally, Section 1.4 provides an overview of the thesis structure.

1.1 Human players and computer players

In computer science and artificial intelligence there have been several attempts to model both human players and computer players (cf. Newell and Simon, 1972). The models applied are usually related to goals ranging from the improvement of a player's strength, via

[1]There are exceptions. The games StarCraft (by *Blizzard*) and Counter-Strike (by *Valve*) are examples of games that have professional players with international rankings.

distinguishing different player types and improving the entertainment value of the game, to identifying user effectiveness in completing tasks using software systems. It is generally acknowledged that such research has a multi-disciplinary nature, involving both computer science and psychology. However, it is rare to find research in this area in which both professional computer scientists and professional psychologists are involved. Therefore, there is much to be gained by a truly multi-disciplinary team (such as ours) working on psychological models of players. In this thesis we have provided definitions of the psychological terms that have been used. These definitions are provided to improve clarity in this multi-disciplinary field. We acknowledge that some researchers or fields of research may have different difinitions for these terms.

Definition 1.1 (Psychological models) Psychological models are models of mental processes that facilitate the prediction of behaviour.

In games research, player modelling is investigated for at least two reasons: (1) to figure out why players behave the way they do, and (2) to improve game content. In parallel to the first reason, personality research investigates personality profiling. Psychologists try to measure differences between people in order to find behaviour patterns that are stable over time. Since stable behaviour patterns help to predict preferences, they enable psychologists to create predictive models. Personality profiling can be seen as the 'real world' equivalent of player modelling.

Definition 1.2 (Player modelling) Player modelling is the practice of creating a model that can be used to predict a player's responses to game content. Player modelling is a technique used to learn a player's tendencies through automatic observation in games (Thue et al., 2007).

Definition 1.3 (Personality profiling) Personality profiling refers to gathering data used for classifying a human's personality.

Both using game research and using self-reports have advantages and missing features. Game research is lacking player models that can be generalised over games; personality research may benefit from games research as an additional method of indirectly and automatically assessing personality that is less susceptible to the problems of self-reports. Both fields can benefit from each other to improve their respective lacunas. This thesis investigates the area that these fields of research share. In particular, we investigate (1) the use of psychological models in creating player models, (2) how to adapt games automatically, and (3) the possibilities of games as a method of automatically generating personality profiles.

For playing games well, humans need (1) skills, (2) experience, and (3) knowledge of the game and the opponents. "Grandmaster" video game players have an excellent mix of these three characteristics at their disposal. The interaction between the three characteristics and video games has so far not been widely examined. We call the three characteristics "human gaming behaviour". So, we investigate whether it is possible to use human gaming behaviour to quantify individual player differences.

We do so by investigating (1) in what respect players differ and (2) how different game properties relate to these differences in players. This explains the title of this thesis: Quantifying individual player differences. If the differences between players are better understood, games can be more effective in achieving their various purposes, whether they are in the field of entertainment, education, health improvement, training, research, or assessment. In our research, we focus exclusively on the player properties in video games; we leave the subject of the other game types to other researchers.

1.2 Problem statement and research questions

The main focus of this thesis is investigating how games can be used to create models of players. Specifically, we want to concentrate on extending existing methods of modelling players with existing theories from the field of psychology. In Chapter 2 we describe the models commonly used in computer science (user, opponent, and player models) and we also give an example of a model from psychology (personality).

Moreover, we want to know whether it is possible to use games as tools that can automatically collect data required to classify users in terms of various psychological models and theories. This focus has led to the following problem statement (PS).

PS: *To what extent are games an appropriate means for measuring the differences between individuals based on psychological theories?*

In this thesis we focus on using games in order to fit psychological models. We limit our scope of investigated models to incongruity and personality models. We do not focus on other models such as models capturing intelligence or attitudes. In order to investigate the problem statement adequately, five research questions have been formulated. The first three research questions involve modelling increasingly complex psychological processes by using games, the final two involve verifying and applying the approaches that we have developed.

Our point of departure is the idea that games may provide an addition to the methodologies currently used in psychology. However, in order to confirm the idea that games can indeed be an addition we need to test whether psychological concepts can be investigated at all with the help of games. We focus on the psychological phenomenon called incongruity. Incongruity is a straightforward process in which one variable influences one result (e.g., incongruity influences emotion). This approach leads us to formulate research question 1 (RQ1) which is examined in Chapter 3.

RQ1: *To what extent are games suitable for measuring incongruity?*

After discussing incongruity, we investigate the effectiveness of using games to model different psychological phenomena. Trait personality theory divides personality into multiple traits. Extraversion is the trait which is validated the most. Extraversion interacts with the situation in which a person is involved and creates a wide range of possible behaviours. We investigate the interaction between extraversion and game behaviour by RQ 2, which is answered in Chapter 4.

RQ2: *To what extent can games be used to measure complex psychological traits such as extraversion?*

Our approach is broadened by focussing on games as a tool for measuring personality. Personality is a system of traits that is potentially useful in the field of player modelling, therefore it merits a more expansive examination. We investigate the potential for automatic player profiling by RQ 3, which is answered in Chapter 5.

RQ3: *To what extent can a data-driven personality profile be created based on game behaviour?*

We have used data-driven methods in building our personality model, psychologists generally use theory-driven methods to construct a personality profile. The effectiveness of this approach is the subject of RQ 4, which is answered in Chapter 6.

RQ4: *To what extent does a theory-driven model explain personality in games?*

After answering RQ 4 we have a better understanding of the way that personality interacts with game behaviour. Because of the number of subjects used in our research, external validation of our models is required. Specifically, the process of establishing player personality needs to be tested in commercial video games (i.e., video games that are commercially available). RQ 5 deals with validating our approach to personality profiling in a commercial video game. RQ 5 is examined in Chapter 7.

RQ5: *To what extent can our models of personality in games be validated in different games?*

Table 1.2 provides an overview of all RQs and the chapters in which they are answered.

1.3 Methodology

The thesis uses mainly experiments to provide answers to the research questions. The common factor between all the experiments is the investigation of human behaviour. Human behaviour refers to both voluntary and involuntary actions performed by humans. The term behaviour covers motions and gestures under the category of physical behaviours, as well as choosing a response on a questionnaire, making choices and decisions in games, and producing verbal responses. In order to avoid confusion, it will be clearly stated for each specific experiment which behaviour is investigated.

We use various methods to record behaviour. Some of these methods are commonly used techniques from the field of psychology. In total, we use four questionnaires and two techniques. The questionnaires are: (1) a general information questionnaire, (2) an incongruity questionnaire, (3) a personality questionnaire (Costa and McCrae, 2008), and (4) an emotional questionnaire (Bradley and Lang, 1994). The techniques are (1) our own technique of automated behaviour logging (van Lankveld et al., 2010), and (2) rated video observation (Back et al., 2009). Table 1.1 contains the abbreviations used for the four questionnaires and two techniques.

Table 1.1: Questionnaires and techniques used.

Abbreviation	Explanation
GQ	general information questionnaire
IQ	incongruity questionnaire
PQ	personality questionnaire
EQ	emotion questionnaire
PBL	player behaviour logging
RVO	rated video observation

Subsection 1.3.1 describes the four questionnaires, viz. GQ, IQ, PQ, and EQ. Subsection 1.3.2 describes the techniques used, viz. PBL and RVO. Subsection 1.3.3 presents the statistical methods used to analyse our data, and Subsection 1.3.4 provides the overview matrix containing the methods used in this thesis.

1.3.1 Questionnaires

One of the commonly used tools in any research involving human participants are questionnaires. Divergent philosophies underlie the different types of questionnaires available. The types we have used are summarised below and described where appropriate.

General questionnaire (GQ)

In Chapters 3 to 8, questionnaires have been used to gather player information. Player information questionnaires are usually presented at the start of an experiment. The questionnaire is meant to collect information that might have an effect on game playing in general. The information collected consists of data such as the age of the player, gender information, education level, computer experience in general, gaming experience, and experience with the game used in the respective research. When this information is collected, items in the statistical analysis can be weighted according to the collected values of the variables in order to examine the influence of these items on the experimental results.

Where participants had to give a numerical value as their answer (e.g., "How much do you agree with the statement that you love to drive cars") Likert-scale items were used (Likert, 1932). Likert-scale items provide a statement for which the participant has to rate how accurately the item describes his attitude regarding a subject (e.g., "Describe how much experience with video games you have"). The ratings range from "no experience at all" to "very experienced". The resolution of the items usually range from 1 to 5 or from 1 to 7. In cases where we wished to exclude the possibility to provide neutral answers the items ranged from 1 to 4 or 1 to 6. In this way the option of choosing a middle answer is removed.

Incongruity questionnaire (IQ)

In incongruity research, our goal is to investigate the relationship of incongruity on player emotion.

Definition 1.4 (Incongruity) Incongruity is defined as the difference in complexity of a context (i.e., the game) and a player's mental model of the game.

We derive the effects of incongruity by measuring the emotions resulting from incongruity between game difficulty and player skill. Therefore, our methodology is to build a game with the ability (1) to measure player skill levels, and (2) to vary the difficulty level in relation to the measured skill level.

Definition 1.5 (Emotion) Emotion is the experience of an internal state or psychophysiological reaction (rather than a cognition).

In order to measure the effects of incongruity, questionnaires were used to assess the emotions predicted by the incongruity theory. The predicted emotions were levels of (1) boredom, (2) frustration, and (3) pleasure when playing a game of a varying difficulty level in relation to the player's skill level. The questionnaire used in assessing these emotions was adapted from van Aart et al. (2008).

This questionnaire consisted of multiple Likert-scale items on which participants provided scores ranging from 1 to 5. The answers on these incongruity items are related to the three emotions described above. The incongruity questionnaire can be found in Appendix A.

Personality questionnaire (PQ)

In order to investigate the differences in player behaviour caused by personality differences a personality questionnaire was used. The personality questionnaire we used is the updated (2008) version of the NEO-PI-R, first developed by Costa and McCrae (1992). This questionnaire is further described and explained in Chapter 2.

Emotion questionnaire (EQ)

Because one of the main goals of our research is to investigate the effects of various game properties on entertainment, we need a measure of entertainment. One of the possible interpretations of entertainment value in games is the amount of positive emotion a game evokes. The term valence is used to describe positive emotions. The emotional model we use is further explained in Chapter 2. We use an emotional questionnaire that measures both positive and negative emotions as well as attention and interest. Our questionnaire is further explained in Chapter 3.

Definition 1.6 (Valence) Valence refers to the amount of emotional attraction or aversion towards a specific object, situation, or event.

1.3.2 Techniques

We use two techniques from the field of psychology to conduct our research. We use PBL and RVO which are based on logging (PBL) and logging and rating (RVO) of behaviour.

Player behaviour logging (PBL)

Player behaviour logging is an adaptation of a commonly used psychological technique for gathering human behaviour known as naturalistic observation (NO). Naturalistic observation is the observation and recording of behaviour in a natural (non-laboratory, non-experimental) setting. The term is used in both human and animal studies (Miller, 1977).

When using NO for psychology or anthropology studies, there are many possibilities in how to record behaviour. In this thesis, the behaviour in a natural setting was "play in a game setting". Many games provide some form of possible logging of events in the game. A main challenge in our research was to decide which game events to log.

Behavioural analysis on humans shows that actions have different meanings in different contexts. Logging the frequency of acts without examining the context can give misleading results. For example, in recording movement one has to classify what the context of the movement is in order to interpret the action properly. Moving toward a character in a game cannot be qualified without knowing who the character is. Moving toward an aggressive looking monster has a meaning different from moving toward an innocent looking small child. The logging of player behaviour is complex and different for each specific outcome. Therefore, we explain the exact logging practice and the reasons behind those choices in the Chapters 4 to 8, respectively.

Rated video observation (RVO)

We use the term rated video observation (RVO) to describe the second technique which is based on NO. In RVO an interview is conducted and recorded. The interview is then rated by several observers who compile a list containing variables related to the research goals. In our case these were personality variables. RVO is further explained in Appendix O.

1.3.3 Statistical techniques

Three statistical techniques have been used during the investigations documented in this thesis. The techniques used are (1) t-test, (2) correlation analysis, and (3) linear regression analysis.

A t-test compares the mean scores of two or more groups for significant differences. In our incongruity research we have investigated the differences in mean scores for boredom, frustration, and pleasure.

Correlation analysis was used to quantify the relationships between individual variables. This was mainly applied in order to clarify the relationships found during the linear regression analysis described below.

Linear regression analysis determines to what degree there are linear correlations for groups of variables. The test corrects for possible differences in a number of observations between variables.

1.3.4 Overview of methodologies used

Table 1.2 shows the various methodological techniques that have been summarised in Table 1.1 (i.e., not the statistical techniques). It shows which techniques have been used in which chapters and where each research question is answered.

Table 1.2: Methodology matrix.

Chapter	Methods used	Questions discussed
1	All	All
2	None	None
3	GQ, IQ	RQ1
4	PQ, PBL	RQ2
5	PQ, PBL	RQ3
6	PQ, PBL	RQ4
7	GQ, PBL, EQ	RQ5
8	None	None
9	None	All

1.4 Thesis structure

Below we provide an overview of the thesis structure. This overview can also be found in Table 1.3. In Chapter 1 we give an introduction to the thesis subject, formulate the problem statement and the research questions, provide insight into the methodologies used, and give an overview of the thesis structure. In Chapter 2 we provide relevant parts of computer science into user modelling, opponent modelling, and player modelling. From psychological theory we provide relevant information as background on entertainment, personality, and emotion. In Chapter 3 we describe our research on incongruity and its relationship to emotion in games and we answer RQ 1. In Chapter 4 we describe our research relating to extraversion and observed game behaviour and we answer RQ 2. In Chapter 5 we present the research on the relationship between the "Big Five" character traits and game behaviour and we answer RQ 3. In Chapter 6 we describe the research on effectively predicting real-world behaviour in the area where personality tests and observed game behaviour are combined and we answer RQ 4. In Chapter 7 we describe the experiment in which we apply the approach introduced in Chapter 5 to a contemporary commercial video game and we answer RQ 5. In Chapter 8 we critically discuss our research findings in a broad perspective. In Chapter 9 we present the conclusions of the thesis and provide advice for future research.

Table 1.3: Thesis structure.

	RQ 1	RQ 2	RQ 3	RQ 4	RQ 5	PS
Chapter 1	✓	✓	✓	✓	✓	✓
Chapter 2						
Chapter 3	✓					
Chapter 4		✓				
Chapter 5			✓			
Chapter 6				✓		
Chapter 7					✓	
Chapter 8						
Chapter 9	✓	✓	✓	✓	✓	✓

Chapter 2

Theoretical background

One of the goals of the social sciences is to model and predict human behaviour. Over the years, this topic has been approached from different angles. In the past decades, computers have become an increasingly popular angle of investigation. In this chapter[1] five perspectives on modelling human behaviour in computers are presented. They are: (1) modelling users, (2) modelling opponents in classical computer games, (3) playing games for entertainment only (we also discuss alternative reasons for playing games), (4) modelling players in games, and (5) psychological models. These perspectives are discussed in Sections 2.1 to 2.5, respectively. In Section 2.6 we summarise the chapter.

2.1 Modelling users

User models (UMs) originate from the field of human computer interaction (HCI) (Fischer, 1999). The purpose of user models is to analyse the way in which an individual interacts with a specific piece of software. After a user model has been created, it can be used to determine what causes the fact that the software is so difficult to use. The software can thus be adapted in order to improve usability.

Definition 2.1 (User model) A user model is an expert system that contains information about a user. The model enables the analysis of the interaction between the user and the software to which the model is applicable.

User models have been investigated in areas ranging from language processing and human computer interaction to intelligent assistants and information retrieval. Chin (1993) describes user models as expert systems containing knowledge about a user. According to Chin, a user model should be able to answer questions about the user, and aid in user-related processes and decisions. The dynamic alteration of software to facilitate effectiveness of use and to promote re-use is an example of such a process.

[1]Partly based on: Bakkes, S. C. J., Spronck, P. H. M., and van Lankveld, G. (2012). Player behavioural modelling for video games. *Entertainment Computing*, 3(3):71–79. I would like to recognise the publisher and to thank my colleagues for their permission to reproduce parts of the article in this thesis.

A typical way of implementing user models is the stereotype approach. A stereotype is the collection of all the relevant characteristics to which the user in the subgroup conforms. Kobsa (1993) gives a fair description of the stereotype approach to user modelling. The approach is divided into the following three tasks: (1) user subgroup identification, i.e., identifying subgroups in a population that possess similar characteristics which are relevant to the application, (2) key characteristic identification, i.e., finding the key characteristics necessary to identify to which subgroup a user belongs, and (3) representation in stereotypes, i.e., forming a stereotype of the subgroup. Yannakakis and Hallam (2007b) apply this approach to games.

One of the possible functions of user modelling is adding adaptiveness to web pages and interfaces (Koch and Rossi, 2002). Web pages can be tweaked to fit more closely to user preference, knowledge, or interest. An example of this form of tweaking is presenting an expert user with more detailed and complex information than a novice user. In general, UMs are applied to increase the ergonomics (e.g., ease of use) of software and to decrease the software learning curve.

Definition 2.2 (Preferences) Preferences are defined as an individual's attitudes.

The term "user model" encapsulates the more focussed terms "opponent modelling" and "player modelling". We discuss these terms below, starting with opponent modelling.

2.2 Modelling opponents

Opponent models are user models applied to opponents in the field of games. They are mainly incorporated within game AI (Bakkes et al., 2012). Below, we first present an overview of game AI, followed by some theory related to opponent modelling.

In the early 1950s, the first empirical research into artificial intelligence was performed. chess and checkers were among the first problems for AI researchers to work on. Shannon (1950) proposed chess as a suitable problem for AI research because chess is sharply defined in both its allowed moves and in its goal. Since chess was a game in which good players were considered to be demonstrating wit (intelligence), it was accepted that if a computer could play chess at a high level it should be concluded that either the computer was intelligent or that the game of chess does not actually require intelligence to play. Either conclusion would be a major revelation for the field of artificial intelligence research.

A simplified explanation for the way that most chess AI works is that it examines all the possible moves at the current point in the game, as well as the possible moves after a move has been made (and so on). An AI can search all the possible modes via incremental depth setting until a favourable outcome (a victory, or, if that is impossible, a draw) for the AI is reached. Van den Herik (1983) gives a more detailed explanation of research on search in chess in the early days. There are many techniques that reduce the number of moves that need to be searched. The techniques are referred to as pruning (Marsland, 1986).

Chess is a complex game; there are usually too many options to consider given the time alotted to make the next move. Because of the time constraint, a chess AI is usually unable to investigate all consecutive moves up to the point where the end condition is reached. In order to deal with this, an estimate of the value of the position reached in the search tree is made by an evaluation function. For chess, this approach works well (see, a.o., van den Herik, 1983). For other games in this category, such as Go, this approach is less suitable and other techniques are used, e.g., Monte Carlo Search (Brügmann, 1993) and Monte Carlo Tree Search (Chaslot et al., 2008)

Using opponent models for chess AI

In the previous section, we have explained that the technique of pruning can make a chess AI more effective. The technique of opponent modelling can also be used to improve the AI effectiveness. Assume that two AI players, named A and B, are playing a game of chess. Player A aims to use opponent modelling. The function of A's opponent modelling is to model player B's decision-making process in order to improve the effectiveness of the moves to be played by player A. Assuming B's evaluation function is not perfect, A could try to find a weakness in B's evaluation function and attempt to exploit that weakness (see Iida et al., 1993a,b; Carmel and Markovitch, 1993; Markovitch and Reger, 2005). In regular tree search, A will play the optimal move from a game theoretical perspective. When A applies opponent modelling, he[2] may prefer a different move because the model predicts that B will respond in a sub-optimal way to the different move. This means that A may obtain an advantage. If B had reacted optimally then A would not have obtained an advantage. It may even happen that A then had to face a disadvantage. Donkers et al. (2003) showed that, by using an incorrect opponent model, a situation may arise in which A is bound to lose without realising that he is in such a special disadvantaged position. So, using an opponent model in chess may have advantages, but should be investigated carefully before being applied to its full extent. In the match Deep Blue - Kasparov (1997), it worked out very well when Deep Blue's operators tweaked Deep Blue's responses specifically to Kasparov's playing style (Campbell, 1999).

Opponent modelling in modern games

While for chess opponent modelling is not a requirement for strong play, for many modern games it is. For poker, opponent modelling is often used for improving playing strength against human opponents (see Billings et al., 1998).

Far more important is the fact that, for modern games, the goal of opponent modelling has shifted, namely from creating strong AI to creating AI that offers an entertaining or appropriate challenge (Iida et al., 1995a; van den Herik et al., 2005). Below we discuss the entertainment that players experience in a game, which is followed by an explanation of how models of players can be used to enhance this entertainment value.

2.3 The entertainment value of games

Our starting point is that most games are played for entertainment. Entertainment can be defined in at least three different ways: (1) as a subjective classification during a specific activity (e.g., the activity is fun/not fun), (2) as a process that evokes positive emotions,

[2]For brevity, we use "he" and "his" whenever "he or she" and "his or her" are meant.

and (3) as a broad attitude or opinion after the experience (e.g., my opinion about strategy games is that they are fun/not fun). Our definition is as follows.

Definition 2.3 (Entertainment) Entertainment is defined as an agreeable pastime.

In Subsection 2.3.1 we look at emotion as explained in psychology. In Subsection 2.3.2 we look at entertainment as subjective experience. In Subsection 2.3.3 we examine the role of emotions in entertainment. In Subsection 2.3.4 we explore the concept of attitudes in relation to entertainment. In Subsection 2.3.5 we present alternative motivations to entertainment for playing games. In Subsection 2.3.6 we present methods to evaluate game experiences.

2.3.1 Emotional experience in games

In order to explain entertainment in games we first provide an overview of human emotion (also known as "affect"). Ryman et al. (1974) investigated self-reports of emotion and noticed 87 emotional terms used to describe emotions. Three examples of these terms are: boredom, frustation, and pleasure. These emotional terms were found to have a degree of overlap. Statistical factor analysis of these terms leads to six basic emotional clusters that accurately collate the underlying terms. The six basic emotional clusters are: happiness, activity, depression, fear, anger, and fatigue.

Russell (1980) determined that the six emotional clusters are not independent. He suggested a so called "circumplex" model that describes two major underlying factors in all emotions: arousal and pleasure. His circumplex model is displayed in Figure 2.1.

In Figure 2.1 the horizontal axis represents pleasure (also known as "valence") (i.e., positive versus negative) and the vertical axis represents arousal (i.e., high versus low). The area outside the circle provides examples where emotional terms fall in the circumplex model. For the example emotions boredom, frustration, and pleasure the states in the model are as follows: boredom is a low negative affect, frustration is a high negative affect, and pleasure is a positive affect (in Figure 2.1, they are indicated by a circle).

2.3.2 Entertainment as subjective experience

The subjective experience of entertainment (specifically games) can be explained in two parts. Both are related to fun. The parts are described below. In passing we note that entertainment in playful activities such as playing video games is often referred to as "fun".

Koster (2004) discusses the first part of what makes up the "concept of an entertaining experience". He presents fun as the process of solving the puzzles that a video game presents. Koster states that video games are basically a series of challenges involving the application of learned skills in the game to novel situations. An example is the following: in the video game SUPER MARIO BROTHERS[3] in the first level the player learns that pressing the B-button performs the jump action. Jumping is a learned skill in the game. The player is then presented with a series of situations in which the key to overcoming the situations is using the learned jump skill. Koster posits that a game is defined as a situation in which

[3]Designed by Shigeru Miyamoto and published by Nintendo in 1987

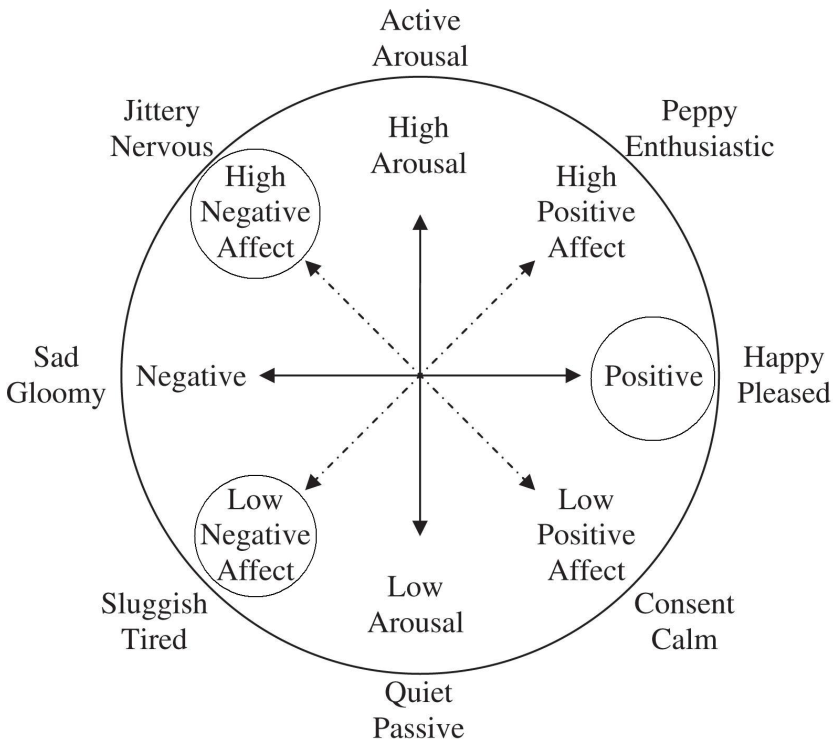

Figure 2.1: The circumplex model of affect.

someone learns to apply skills effectively in order to overcome obstacles. Moreover, single-player games usually feature unlockable abilities or powers that require significant learning to master.

Definition 2.4 (Entertaining experience) An entertaining experience is defined as an enjoyable pastime.

The second part of fun in games is that games require what is referred to as "a well shaped difficulty curve" (Aponte et al., 2011). This is a concept that is also given attention in the academic community (see Chapter 3 of this thesis, and Rauterberg, 1995). The definition of "a well shaped difficulty curve" is vague. An often found definition is that the difficulty of a game should neither be too low nor too high, but should fit a player's skill level. In many games, players can manually alter the difficulty curve of a game by adjusting the difficulty level at the start of the game and sometimes even during the game. The earliest example of a game featuring variable difficulty setting was Speed Race[4].

[4]Designed by Tomohiro Nishikado and published by Taito in 1974

Definition 2.5 (Well shaped difficulty curve) A well shaped difficulty curve is defined as a progression of game difficulty during play that does not evoke negative emotions in the player.

Currently, commercial games usually provide a manual way of setting difficulty at the start of a new game. This method results in an inadequate difficulty setting if the player makes an unsuitable choice or if his skill improves during play. For example, the commercial game MAX PAYNE[5] features what the developers refer to as "dynamic difficulty adjustment" (DDA). The DDA monitors the amount of damage received, and adjusts the player's auto-aim assistance and the strength of the enemies. This approach is easily recognised by the players, and breaks the flow of the game (some phenomena that break flow were already identified in 1988 by Csikszentmihalyi (1988)). Recognising the mechanism may lead to players taking extra damage on purpose in order to decrease the game's difficulty.

Definition 2.6 (Dynamic difficulty adjustment) Dynamic difficulty adjustment is defined as a method of automatically altering the game difficulty to suit the player's skill level.

Computer science researchers have investigated methods to measure the entertainment value of a game (Iida et al., 1995b; Yannakakis and Hallam, 2007a,b; Beume et al., 2008), and sometimes even to adapt the game automatically in order to increase entertainment (Hunicke and Chapman, 2004; Spronck et al., 2004). Yannakakis and Hallam (2008) describe two ways of optimising player enjoyment, namely implicit and explicit. In implicit optimisation, machine learning techniques, such as reinforcement learning, genetic algorithms, probabilistic models, and dynamic scripting, are used for optimisation. They also mention user modelling techniques used in interactive narration. In explicit optimisation they describe adaptive learning mechanisms used to optimise what they call "user verified ad-hoc entertainment".

Definition 2.7 (Implicit optimisation) Implicit optimisation is defined as a form of optimisation that is automatic and that proceeds without confirmation by the player.

Definition 2.8 (Explicit optimisation) Explicit optimisation is defined as a form of optimisation that can be verified by the player.

2.3.3 Entertainment and the experience of positive emotion

The academic concept of fun in video games was pioneered by Csikszentmihalyi (1989). Csikszentmihalyi is a researcher in the domain of positive psychology who investigated the features of intrinsically rewarding experiences. Intrinsically rewarding is a term indicating activities that are rewarding for their own sake. For an intrinsically rewarding activity, no reward from outside is needed to motivate a person to perform the activity. According to Csikszentmihalyi (1989), flow is an emotional state belonging to intrinsically rewarding activities. The flow state resides between the states of boredom and anxiety (see Figure 2.2). In the figure the vertical arrow represents the effects of an increasing challenge while the horizontal arrow represents the effects of an increase in skill. Neither skill nor challenge

[5] Developed by Remedy Entertainment and published (in the EU) by 3D Realms in 2001

are exactly determined, therefore there are no units of measurement displayed. Skill and challenge are expressed differently for different activities.

Flow is experienced by performing tasks that feature just the right amount of complexity or challenge. According to Csikszentmihalyi, in order to reach the flow state, both a high challenge for the task and a high skill for the person completing the task are required. Test participants of Csikszentmihalyi reported the feeling of flow to be akin to feeling active, alert, concentrated, happy, satisfied, and creative. Tests for of cheerfulness and sociability were not associated with flow emotions.

Definition 2.9 (Intrinsically rewarding) Intrinsically rewarding is a definition for activities that are rewarding for their own sake.

Definition 2.10 (Flow) Flow is defined as an emotional state of high challenge and high skill that is perceived as enjoyable.

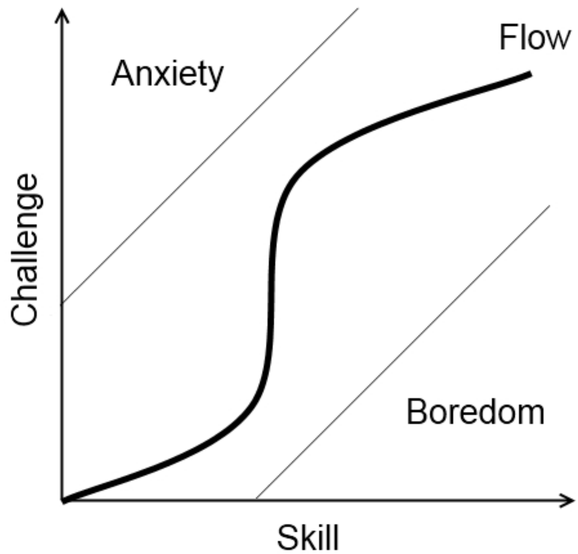

Figure 2.2: Schematic representation of flow.

The concept of flow has been adapted for use in computer games. Sweetser and Wyeth (2005) state that adapting flow theory for games leads to eight elements by which games can be evaluated: concentration, challenge, skills, control, clear goals, feedback, immersion, and social interaction. They evaluate two games based on these eight elements. These games are Warcraft 3[6] and Lords of Everquest[7]. They are in the same genre and have

[6]Designed and published by Blizzard Entertainment in 2002

[7]Developed by Rapid Eye Entertainment and published by Sony Online Entertainment in 2003

roughly the same content. However, the first game received generally positive reviews while the second game received generally negative reviews[8]. Both games were evaluated based on the flow criteria and the flow related performance was found to match the review scores. However, this research can be criticised for the fact that Sweetser and Wyeth (2005) only performed a post-hoc analysis of the correlation between these games and the review scores: they did not test on the ability to predict review scores for games that had not been reviewed yet. It is also notable that they did not correlate their research with experienced emotions during play for the investigated games.

2.3.4 The role of attitudes in entertainment

Games companies often conduct market analyses in order to increase sales of their products. Game contents are scrutinized carefully in order to see which elements of games are positively received and which are not. When companies make sequels they often include much loved thematic elements and characters from the previous game in the series. Moreover, additional research is invested into finding out which elements and characters are popular and which are not. This form of analysis is referred to as "game metric analysis" or "usability-testing" (Tychsen, 2008). It is usually combined with user experience surveys. We note that a player's attitude toward a game after playing might not correlate with his experiences during the game.

The attitudes of people towards a game is an important topic for the serious games community. One of the reasons often cited for using games as learning tools is that children are not motivated to go to school but they are motivated to play video games. In this line of reasoning video games are considered a preferred medium for teaching (Prensky, 2003).

Hsu and Lu (2004) have investigated which factors contribute to the intentions of users to play online games. They adapt the technology acceptance model (TAM) to games. The TAM is based on the theory of reasoned action (TRA) which states that an individual's belief influences the attitude. In turn, the attitude shapes behavioural intention (BI). In the case of games: beliefs about games shape our appreciation of games (attitudes), which shape our intention to play games (or a specific game). For example, if a player is bored playing real-time strategy games, he might adopt the attitude that real-time strategy games are boring. This, in turn, could decrease his intention to play real-time strategy games. Hsu and Lu (2004) conclude that for games there are two factors which influence attitudes about games: (1) flow experience (see 2.3.3) and (2) social norms (i.e., peer pressure).

2.3.5 Alternative reasons to play games

Contrary to the popular opinion, people might play games for other reasons than the feeling of fun or entertainment. Csikszentmihalyi and Csikszentmihalyi (1989) hint at this possibility. They state that people commonly feel more flow at work than during leisure time. They speculate that a possible reason is that flow experiences can be exhausting and people might need to rest during their leisure time. By extension, games might be used as a form of leisure to relax oneself at the end of a day of hard work. If this is the motivation for playing a game, high complexity might not be a preferred attribute of the game. Even though in

[8] http://www.metacritic.com/game/pc/lords-of-everquest

this example, video games would not be played for experiencing a flow sensation, it could be argued that the games are still played for relaxation.

There are three alternative reasons to fun for playing games (see Susi et al., 2007; Wong et al., 2007; Tejeiro Salguero and Morán, 2002; Grüsser et al., 2007; Chou and Ting, 2003; Schull, 2002): (A) serious games are played in order to facilitate learning or training, (B) games might be played because the player is addicted to gaming, and (C) games might be used as a mechanism for coping with problematic situations in other parts of life.

A: Learning or training

Serious games are used in many different areas (Susi et al., 2007). Examples are: military, government, education, corporate, and healthcare. The function of serious games is to educate or train students and professionals. A precursor to serious games is edutainment, which, according to Susi et al. (2007), did not produce successful results. Users of edutainment reported feelings of boredom and monotony. Training and learning proved to be less successful than when using conventional forms of training and learning. There is also some controversy about serious games themselves. Susi et al. claim that the evidence for the supposed learning benefits in serious games is scarce because extensive experiments are lacking. Wong et al. (2007) provide evidence that serious games are more suitable than text and hypertext in transferring knowledge about molecules in the human body.

B: Addiction in games

Tejeiro Salguero and Morán (2002) and Grüsser et al. (2007) investigate the relationship between gaming, substance dependence, and gambling addiction. They show that excessive use of games shares symptoms with the type of addiction known as a dependence syndrome, and they show that roughly 12% of gamers suffer from these symptoms. Chou and Ting (2003) also show that gamers can suffer from addiction to gaming. Gaming can become an obsession, when a person continues to play the game even though he feels it is no longer in his best interest. Chou states that addiction is more likely to occur if a player has experienced flow in a game. He demonstrates that when repetition of gaming triggers a flow state the chances of developing an addiction are greatly increased.

C: Coping mechanism

A possible explanation for gaming addiction is coping behaviour, as found in gambling addiction (gambling and gaming are closely related). Schull (2002) relates gambling on video slot machines to escaping emotionally demanding home situations. She describes how house wives seek means to relieve their anxiety and tensions at home and find the solution in gambling. In the same vein, players might use games to avoid thinking about traumatic or emotionally taxing situations.

2.3.6 Methods to evaluate entertainment

A common method of sampling attitudes, opinions, and experiences in the social sciences is the survey (or questionnaire) (Goodman, 1997). Csikszentmihalyi and Hunter (2003) have developed a survey to measure experiences of personal happiness. They state that happiness is a combination of five factors: (1) genetic determinants, (2) macro-social conditions, (3) random events, (4) environment, and (5) personality. Cziksentmihalyi's survey allows for the investigation of environmental changes on happiness. More specifically, it allows for measurement of happiness in various situations (called "environments" by Czikszentmihalyi). The method is called Experience Sampling Method (ESM). ESM consists of a questionnaire that should be filled out whenever a pager gives off a signal. The questionnaire asked what the participant was doing at the precise moment of the signal, as well as gave multiple choice questions and scales, asking for the participant's feelings at that time. For definitions of both happiness and ESM, see Csikszentmihalyi and Hunter (2003).

2.4 Modelling players

Entertainment is a subjective experience. In order to be appropriately entertaining for a specific player, the game must make a connection with who the player is. While some game developers are seldom concerned with this topic – they generally assume that a large group of players will find their game entertaining straight out of the box – creating a successful model of a player has recently been the subject of game research.

Player modelling concerns generating models of player behaviour and exploiting the models in actual play. Considering the increasing complexity of state-of-the-art video games (Rabin, 2008), player models are sorely needed for (1) predicting the player accurately and (2) adapting to the player. In general, a player model is an abstracted description of a player in a game environment. Specifically for the context of behavioural modelling, a player model is an abstracted description of a player's behaviour in a game environment. In general, it concerns only the behaviour of human players. However, player modelling techniques can also be applied to the behaviour of AI controlled characters.

Definition 2.11 (Player modelling) Player modelling is the creation of a player model. Player models are described below.

The general goal of player behavioural modelling is to steer the game towards a predictably high player satisfaction (van den Herik et al., 2005) on the basis of modelled behaviour of the human player. Moreover, next to being useful for entertainment augmentation, player models may be useful for simulation purposes (e.g., simulating stories or evaluating game maps), game design purposes (e.g., testing whether the map leads to the game play as envisioned by the designers), and serious game applications such as education (e.g., tailoring the game to a players model for reaching particular learning objectives) or health (e.g., personalising games for rehabilitation of elderly patients).

Fürnkranz (2007) states that player behavioural modelling is of increasing importance in modern video games. The main reason is that player behavioural modelling is almost a necessity when the purpose of AI is entertaining the human player rather than defeating the human players (see also van den Herik et al., 2005). A challenge for such player modelling in video games is that models of the player have to be established (1) in game environments that generally are realistic and relatively complex, (2) with typically little time for observation, and (3) often with only partial observability of the environment. The online creation of player models, as well as the (optional) classification of the player into previously established models, is a task that has to be performed in real time, while other computations, such as rendering the game graphics, are performed simultaneously. Researchers estimate that generally only twenty per cent of all computing resources are available to the game AI (Millington, 2006). Of this twenty per cent, a large portion will be spent on rudimentary AI behaviour, such as manoeuvering game characters within the game environment. This implies that only computationally inexpensive approaches to player modelling are suitable for incorporation in the game AI.

Player models (PMs) are essentially models of the current state of the player. The state may include the emotions of the player, his preferences, and his goals. Two functions of PMs are (1) to increase the entertainment value of the game and (2) to decrease the amount of player frustration concerning unwanted behaviour of game characters.

Definition 2.12 (Player model) A player model is a model that contains state information about a player of a video game. Player models may be used to alter game content and game behaviour for the purpose of entertainment.

Player modelling has been used as a basis for making the behaviour of AI-controlled non-player characters (NPCs) more human-like. PMs are also used to adapt the content of the game dynamically. Two examples of player modelling attempting to enhance the entertainment of games are the research by Thue et al. (2007) and by El-Nasr (2007), in which PMs are used to adapt the story and action in the game in order to fit the player's preferences. Below we discuss modelling player actions (2.4.1), skill-related player differences (2.4.2), and player types (2.4.3).

2.4.1 Modelling player actions

A straightforward way to implement player modelling is by modelling the actions that a player executes. Such an action model consists of a list of game states, each combined with one or more player actions, and a likelihood value that the player will undertake that action in the state. A perfect action model predicts exactly one action for each possible game state with 100% accuracy.

Definition 2.13 (Action model) An action model is defined as a specialised player model that contains data on the actions which a player may perform for the various game states.

Opponent models, as used in classic board games (see 2.3) are typically action models, as they predict the moves that the opponent is expected to make. Note that actually all tree-search techniques use action models. As by default they sometimes use the computers' own evaluation function to predict opponent moves; then, the opponent model used is actually the computer itself.

Definition 2.14 (Opponent model) An opponent model is a model of an opponent's set of strategies that allows for increasing the effectiveness of play against that opponent.

The first explorations of the kind of player models that can be used in video games was performed in RoboSoccer. RoboSoccer as a research environment is comparable to video games such as sports games and first-person shooters. The models were predominantly action models, which specifically predict what kind of actions the opponent bots are going to take. For example, Ledezma et al. (2005, 2009) used classification techniques to build action models of members of the champion team of the 2001 RoboCup edition.

A straightforward technique that has been proposed for building action models for video games is sequential prediction (Mommersteeg, 2002), specifically by the use of N-grams (Laramee, 2002). N-grams are sequences of choices, i.e., moves or actions. It is assumed that action sequences that have been observed in the past can be used to predict a future action. For instance, if it has been observed that when action A_1 is executed twice in a row, it is followed 75% of the time by action A_2, the prediction would be that there is a 75% likelihood of the next action being A_2 if the previous two observed actions were both A_1. In general, the more actions in the past are observed, the better the N-grams will function. A problem with N-grams is, however, that they are only based on action sequences, while disregarding other state parameters. Therefore they mainly work for games in which the prediction of move sequences is key to game-play, such as fighting games.

In many video games the number of low-level actions is so large that it is hard to predict which one the opponent will take. However, actions might be predicted on a higher level where the number of possible actions is manageable. Work by Butler and Demiris (2010) uses an approach inspired by the Theory of Mind, in which they predict the selection of a target of a team of units in an RTS game, by mapping the team's movement to A* paths which lead to the respective targets.

An advantage of action models is that they are easy to employ by an AI. If it is known which action the opponent is going to take, it is easy to block the action or avoid confrontation, if desired. However, there are two drawbacks.

The first drawback is that states in video games typically encompass a large number of parameters, and the number of different actions is usually also large. This leads to an unmanageable state-action space. Moreover, in most games the state information is incomplete. The consequence is that for action models to be learned efficiently, state information must be restricted to a few simplified features, which are usually insufficient for building an acceptable action model except for rather straightforward games.

The second drawback is that action models do not generalise well, as the reason why a player takes an action is not part of the model. For example, assume that a human player controls a fighter character in a role-playing game, and an action model is determined for his behaviour. When later the human player controls a wizard character, with a different list of possible actions, the previously learned action model has become useless, even though the player might still employ a similar playing style. Action models therefore can be useful in relatively low complexity games, but do not scale well to more complex games.

2.4.2 Skill-related player differences

One of the most obvious differences between players is the difference in playing skill. Players have different effectiveness in reaching game goals. This difference is sometimes referred to as a difference in skill. Skill is an accumulation of different factors. Skill is influenced by the amount of practice a player has had and by the player's natural ability of exploiting this practice (sometimes referred to as intelligence). Skill is also influenced by the player's natural reflexes and speed of thought in a given game situation (also sometimes referred to as intelligence). These skill-related factors are not widely examined in the field of video games.

Definition 2.15 (Skill) Skill is defined as an individual's effectiveness in reaching goals.

Definition 2.16 (Skillful) Skillful is a connotation signifying a high quality with which a task is performed.

Most games have a measure of skill. The most common measures of skill are ratings and scores. An example of a rating is the Elo rating in chess (Elo and Sloan, 2008). Similar ratings are also in use in other classical board games. Examples of scores are the high score ratings at the end of an arcade game and the online player ranks in the multi-player part of first-person shooters. We note that the player level in massively multi-player online role-playing games (MMORPGs) is not an accurate measure of skill. The level of MMORPG characters increases more reliably with time played than with the greater skill of a player. Game adaptation based on skill will be further explained in Chapter 3.

2.4.3 Player types

Player types have received specific attention in the player modelling field (Bartle, 1996; Drachen et al., 2009). Player types usually classify gamers based on behaviour or preferences. Bartle (1996) uses four types of players do describe dynamics of change in a MUD game community. This test has been adapted to be used as an online test of gamer personality[9] and has been taken over 760 thousand times. While, from a testing methodology standpoint this test has it's flaws, it is undeniable that a sample size of over 760.000 has the potential to show stable psychological traits in behaviour. Drachen et al. (2009) identifies player types by using emergent self-organising maps.

Definition 2.17 (Player type) A player type is a method of classifying players into types based on their playing behaviour or their preferences.

For many games a limited number of player types can be distinguished, each with a predisposition for specific action choices. An action model of a particular player can then be defined as a series of weights for each of the possible player types, and the predicted choice of action can be determined as a weighted voting by all the types. This is the basis behind the strongest player models for Texas Hold'em Poker (Billings, 2006), but is also used for other games, such as Guess It (Lockett et al., 2007).

The previous examples show that player types are avalid approach to player modelling. Types have been used in early personality psychology. However, most personality researchers now prefer a factor approach to personality theory.

2.5 Psychological modelling

The most obvious differences between people are age, gender, and race. However, there are many more factors that make people unique. People have unique genotypes, phenotypes,

[9]http://www.gamerdna.com/quizzes/bartle-test-of-gamer-psychology

histories, and environments. For the topic of video games physical differences are far less important than mental differences (with the possible exception of physical motor skills). All these differences may have an impact on player behaviour.

In the playing of games there are two types of behaviour: (1) behaviour related to the attainment of the game goals and (2) all other behaviour. Other behaviour includes expression of individuality, for instance, in the appearance of game avatars. The appearance of game avatars usually has no effect on the effectiveness with which game goals can be reached. A second important behaviour is the choice of a game and of the game goals inside that game.

In this section we discuss the differences between players that are unique to the player but generalisable across games. This means that we assume that player differences are properties of the player that will manifest themselves across games with a specific expression inside each game. In Subsection 2.5.1 we discuss the differences between players not related to skill, i.e., personality theory. In Subsection 2.5.2 we discuss the origins of personality. In Subsection 2.5.3 we discuss contemporary views on personality theory.

2.5.1 Personality theory

Players differ in personality. Personality is defined as a set of stable factors that influence a person's behaviour in some way. Psychology is concerned with classifying, describing, predicting, and manipulating human behaviour. Human behaviour is an expression of human cognitive, emotional, and autonomous processes, such as reflexes and the physiology of the body. It is influenced by experiences gained in the course of time. Human behaviour comes in many forms like speech, communication, and facial expressions. In this sense answering a questionnaire can also be considered to be human behaviour.

Definition 2.18 (Personality) Personality is defined as the stable pattern of variation in individual acting, thinking, and experiencing.

Personality influences many human behaviours, some even say it influences most human behaviours. For our research, personality theory is an important fundament. There are two reasons for this: (1) since it influences so many behaviours and remains stable across many different situations, personality is a prime candidate to base game adaptation on; (2) personality traits are hard to measure, currently the bulk of personality research focusses on questionnaires. Games might provide an alternative to questionnaires for researchers who are interested in measuring personality.

Personality theory describes a person's stable attributes, emotional patterns, and intentions. These attributes are patterns of behaviour and responses caused by factors in the person. Currently, personality traits are viewed as psychological constructs that describe major categories of behaviour. There are various fields of personality theory. In this thesis we will be focussing on the five factor model (FFM) of personality, also known as the "Big-Five". The five factor model is briefly described in Subsection 2.5.3.

Definition 2.19 (Personality theory) Personality is a class of psychological theory that is concerned with personality.

2.5.2 The origin of personality theory

Comparisons between people are commonly based on traits (Gosling et al., 1998). The earliest known personality descriptions were suggested by ancient philosophers. They first explored personality through observation and reasoning. They tried to understand illness, emotional suffering, and behaviour (Magnavita, 2002). Thinking about personality followed a logical rather than empirical line of thought.

In the 19th century psychiatry explored personality in an attempt to cure mental illness. Freud and Jung were amongst the first to examine properties of the mind in order to diagnose dysfunctional behaviour (Glover, 1991). Freud's ideas were based on personal philosophies, while Jung required empirical evidence and facts to support his theories. Jung's ideas are at the basis of modern psychology (Smith, 1977).

If a psychological theory is empirically validated and the model is standardised it can be used to compare individuals to groups of people. William Wundt started the empirical validations of personality by using experimentation. Wundt laid the basis for modern experimental research methodology, and investigated various domains of psychology including consciousness, perceptions, sensations, and feelings (Magnavita, 2002). These accomplishments lead directly to the domain of psychological profiling.

In 1936 Allport and Odbert (1936) identified nearly 18,000 words in the English dictionary that were used for the description of personality. In the following 75 years, dozens of personality questionnaires have been developed that are based on these descriptive dictionary words. This type of approach is called the *lexicographic method*.

2.5.3 The five factor model

At the start of the 20th century personality theory was seen as a chaotic and unstructured field. Personality was being investigated at different levels of abstraction and from different perspectives. In order to give structure to the field of personality research, a descriptive model was needed. One taxonomy was found in which the entire field could be represented: the five factor model of personality. We briefly describe the development of the five factor model. Thurstone (1934) was the first to suggest a system of five domains. Thereafter, several other researchers found evidence for a system of five factors as well. This marked the start of the five factor model (Wiggins, 1996). Even though there is some discussion on the correct names, the five factors are most commonly named (1) Openness to new experiences, (2) Conscientiousness, (3) Extraversion, (4) Agreeableness, and (5) Neuroticism (which can be abbreviated as OCEAN). The five factor model was based on the terms people use to describe each others' stable attributes. The model divides personality into five domains by which a description of someone's personality can be given. The model was designed by analysing the natural language terms people use to describe one another (John and Srivastava, 1999).

2.5.4 The NEO-PI-R

The five factor model was independently confirmed in several studies but received near fatal criticism in the 1970s and 1980s. Mischel (1972) criticised the trait approach in general and disputed the reliability of five factor research up to that time. Costa and McCrae (1992) also provided criticism but then provided a more reliable instrument as the solution to the criticisms: a new personality questionnaire named NEO-PI, the first robust tool for measuring the five factor model. NEO-PI is an abbreviation for Neuroticism, Extraversion, and Openness to experience Personality Inventory. The earliest versions of the NEO-PI measure only three personality traits (neuroticism, extraversion, and openness); in the following years two other traits were added (agreeableness and conscientiousness).

The NEO-PI divides every trait into six facets. These facets provide a detailed specification of the contents of each domain (Costa and McCrae, 1992). The facets were designed to be supported by existing literature. They were meant to be similar in breadth and should represent "maximally distinct" aspects of each domain.

A more modern test, the NEO-PI-R (the 'R' standing for 'revised'), is now considered a reliable and valid test for personality. It contains 240 items measuring the five domains and their facets. It has been thoroughly tested (Costa and McCrae, 1992), and is widely accepted as the standard model of personality structure (Goldberg, 1993).

The domains of the five factor model as labelled by Costa and McCrae and tested by the NEO-PI-R are: Openness to Experience, Conscientiousness, Extraversion, Agreeableness, and Neuroticism (OCEAN) (McCrae and Costa, 1997). Trait scores follow a normal distribution. We will give a description per trait in terms of behaviour that can be seen in natural human settings. The precise description of a trait tends to vary slightly among researchers. Here we adhere to the common descriptions. The NEO-PI was meant to replace earlier, suboptimal tests measuring the five factor model (Costa and McCrae, 2008).

- Openness: The interest in novel stimuli. A high score is typically accompanied by curiosity and willingness to deviate from social conventions.

- Conscientiousness: The propensity to adhere to rules, both social and personal. This trait is also tied to the ability to restrain oneself and the ability to stick to a plan during periods of stress and difficulty.

- Extraversion: High scorers seek excitement and positive stimuli. This often leads to individuals seeking the company of others and seeking exhilarating situations like high speed driving, roller coasters, and other high adrenaline activities.

- Agreeableness: Explained as compliance, willingness to cooperate, and friendliness. Low scorers tend to follow their own needs over those of others. High scorers are seen as empathic.

- Neuroticism: This trait is connected to fluctuating and negative emotions such as anger and fear (see Figure 2.1). High scorers are more likely to check situations for safety. There is also a relationship to shyness and social anxiety.

The NEO-PI-R measures 6 facets per trait. For this reason the 240 questions are needed. There is a more concise version called the NEO-FFI which measures only the traits and not the 6 facets per trait resulting in a less exhaustive but shorter version of the test (Hoekstra et al., 2007).

2.5.5 Contemporary methods for measuring personality

Currently, the most commonly used methods of measuring personality are the following tests/procedures.

1. Written tests: Written tests are usually lists of statements describing personal preference and behaviour. In such a test, subjects are invited to rate to what degree the statements describe them correctly. Based on these ratings, a personality profile is computed.

2. Verbal tests: Verbal tests are interviews in which a psychologist asks a subject questions about his preferences. Then he composes a personality profile based on the subject's answers.

3. Observational studies: In observational studies a trained observer analyses a subject directly or scans videos of a subject, and composes a personality profile based on the observed behaviour.

The three methods can be applied during personality testing, they suffer from four drawbacks.

1. Written tests and verbal tests are based on the assumption that a subject's reports are truthful and comprehensive. When a self-report is inaccurate or untruthful the validity of the test is decreased. It has been shown that subjects are unable to report

accurately on their own habits. Gross and Niman (1975) have pointed out that self-report data have little correlation to actual behaviour frequencies. Still, questionnaires are frequently used because rating personality requires many samples of data across a wide range of situations.

2. The need for extensive and wide data gathering makes interviews and observation time consuming and expensive.

3. Questionnaires provide a reasonable alternative to interviewing and observing but their advantage in time requirements comes with a decrease in reliability (Kolar et al., 1996).

4. Observational studies are considered to be more reliable and more objective than self-reports (Arney, 2004). They do not suffer from inaccurate subject reports. However, these studies suffer from high cost and high effort in data collection. Gathering sufficient data through observational studies to form an adequate model of personality may take years of work and may involve numerous observations on numerous subjects (McCrae and Costa, 2003). The fourth drawback may be overcome by the introduction of the automated personality profiling techniques described in Chapter 4 through Chapter 6.

Personality tests in which a subject knows that his personality is tested are called explicit tests. Explicit personality tests are vulnerable to socially desirable behaviour. People tend to act more socially favourable when they feel they are being evaluated or assessed. They do so by presenting themselves in a more accepted fashion. An example is: people pretending to be more conscientious than they really are (cf. Fisher, 1993).

2.6 Chapter summary

In this chapter we presented the basis for the research discussed in this thesis: player modelling. We explained the history of player modelling, starting with user modelling and opponent modelling. We then explained that the goal of many modern games is to provide entertainment, which requires knowledge of a player's characteristics (as captured in a player model) rather than knowledge of this behaviour as an opponent (captured in an opponent model). As a basis for a psychologically valid player model we introduced personality theory, in particular, the five factor model of personality, as measured by the NEO-PI-R.

In the next chapters, we present experiments to demonstrate the power and applicability of player models, starting with a model based on incongruity theory in Chapter 3.

Chapter 3

Incongruity in games

In this chapter[1] we examine RQ1: *To what extent are games suitable for measuring incongruity?* We investigate incongruity and its relationship to the emotions of players at varying levels of game difficulty. Incongruity is the difference between the complexity of the game and the complexity of the mental model that the player has of the game. We construct a game that measures the effectiveness of players in defeating three types of enemies. We apply these measurements to keep the difficulty level of the game constant with respect to the experience level of the player. The experience level of the player is considered to be a measure of the complexity of his mental model. We perform an experiment in which we observe the reported emotions for different levels of incongruity in the game. Because we are monitoring incongruity, we are creating a skill-based player mode as a result.

The outline of the chapter is as follows. In section 3.1 we explain the theory of incongruity. The game GLOVE for our incongruity experiment is described in Section 3.2. The experimental setup is discussed in Section 3.3. The results are presented in Section 3.4 and discussed in Section 3.5. Section 3.6 concludes and answers RQ1.

3.1 Incongruity

People continuously form mental models about the world in which they live. These models allow them to estimate how the world is going to react to different types of interaction and how the world will change over time. Incongruity theory attempts to explain the emotions that arise during play as a consequence of the difference between the complexity of the mental models and the world.

Definition 3.1 (Incongruity theory) The incongruity theory attempts to explain the three variations in emotions caused by incongruity by the assignment of three emotional states: (1) boredom, (2) frustration, and (3) pleasure.

[1]Based on: van Lankveld, G., Spronck, P. H. M., van den Herik, H. J., and Rauterberg, M. (2010). Incongruity-based adaptive game balancing. In *12th International Conference, ACG 2009*, pages 208–221. Springer-Verlag, Heidelberg, Germany. I would like to recognize the publisher and to thank my colleagues for their permission to reproduce parts of the article in this thesis.

In order to clarify the incongruity theory we explain it step-by-step in the subsections below. In Subsection 3.1.1 we describe the terminology used in incongruity theory, in Subsection 3.1.2 we describe the predicted relationship between incongruity and experienced emotions during play, in Subsection 3.1.3 we discuss the origins of incongruity theory, and in Subsection 3.1.4 we present the contemporary uses of incongruity theory.

3.1.1 Incongruity theory

Incongruity theory predicts the emotions that arise due to the interaction between player skill and game difficulty. In incongruity theory the term "context" is used to refer to the world, or part of the world. In our case, this is the game or the game's areas. The term "mental" is used to refer to the mental model belonging to a person who interacts with the context (i.e., who plays a game). The term "context complexity" is used to describe the complexity of the context. The term "mental complexity" is meant to describe the complexity of a mental model that a person has of the context. The term"incongruity" is used to describe the difference in complexity between the context and the mental model. When the context complexity is higher than the mental complexity we speak of "positive incongruity" (Figure 3.1 A). When the difference between the context complexity and the mental complexity is small we speak of "no or low incongruity" (Figure 3.1 B). When the mental complexity is higher than the context complexity we speak of "negative incongruity" (Figure 3.1 C). Mental models tend to develop (i.e., increase in complexity) when confronted with a situation with a positive incongruity: this is "learning".

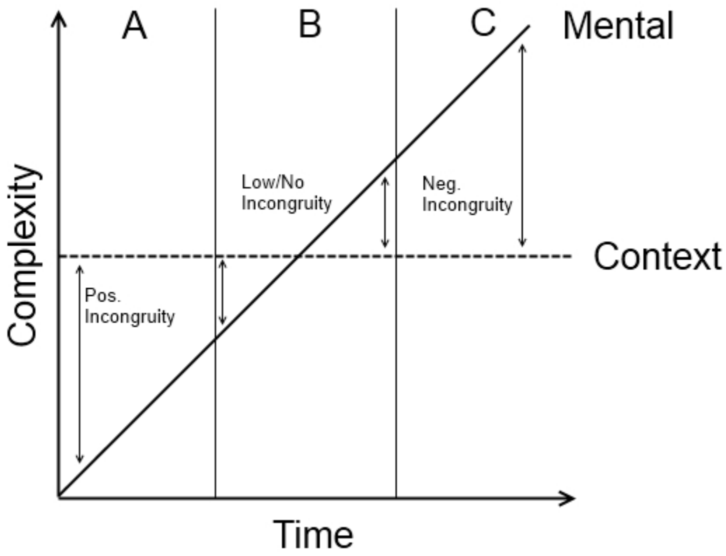

Figure 3.1: Schematic representation of incongruity.

To clarify the figure we provide an example of incongruity in chess playing. When a novice player starts to learn how to play chess, he could try to practice playing against the CHESS CHAMPION AI (this is a low level AI from the 1980s). When the novice starts playing there is high positive incongruity; the chess AI is much more complex than the (low) mental complexity of the novice player. As the player learns more about chess his mental complexity rises and incongruity decreases. Eventually, the player reaches the level of the AI and there is no more incongruity; the mental complexity of the player is at a point equal to the context complexity.

If the AI plays at a stable difficulty level (a stable level of complexity) and the player continues to improve, then eventually the player's mental complexity will surpass the context complexity. This results in negative incongruity. Adjusting the level of the CHESS CHAMPION AI to a higher level may then raise its complexity level above the player's mental complexity and result in positive incongruity once again.

Because mental complexity is part of the human mind there is usually no direct way to measure mental complexity in games. One possibility is to measure the complexity of the player's behaviour in a game and infer the mental complexity from that. Rauterberg (1995) stated that:

> *"If the cognitive structure is too simple, then the concrete task solving process must be carried out with many heuristics or trial and error strategies"*

3.1.2 Incongruity and emotions

According to the incongruity theory, the difference between the mental complexity and the context complexity may give rise to three emotions: boredom, frustration, and pleasure. Boredom is a feeling of reduced interest, which arises with high negative incongruity (a player with this mental model has no difficulty in winning the game). According to Rauterberg (1995), in situations of high negative incongruity, people start to look for new a stimulation. Frustration is a feeling of annoyance or anger, which arises with high positive incongruity (a player with this mental model has no chance to win the game). According to Rauterberg (1995), in situations of large negative incongruity, people give up playing. Pleasure is a feeling of entertainment, it arises when context complexity is roughly equal to or slightly higher than mental complexity. According to Rauterberg (1995), in situations of low incongruity, people enjoy themselves and continue playing. For a more extensive explanation of emotions we refer to Chapter 2.

Definition 3.2 (Boredom) Boredom is defined as a feeling of reduced interest.

Definition 3.3 (Frustration) Frustration is defined as a feeling of annoyance or anger.

Definition 3.4 (Pleasure) Pleasure is defined as a feeling of entertainment or enjoyment.

For the CHESS CHAMPION AI example the player could feel frustration when he starts playing against the AI if the incongruity is highly positive. As positive incongruity decreases by the player's learning, the player starts to feel pleasure from playing. When negative incongruity develops, the player starts feeling bored. Learning is stimulated in situations where incongruity is positive: it raises the mental complexity. Thus, learning can bring players from a large positive incongruity via low or no incongruity to negative incongruity.

3.1.3 Origins of incongruity theory

The original concept of incongruity was proposed by Hunt (1963). Rauterberg (1994) produced an adaptation of Hunt's concept to use it in the field of human-computer interaction. This adaptation is the incongruity theory described above. Rauterberg also bases the properties of his incongruity theory on a paper by Smith (1981) which summarizes research on the relationship between boredom, frustration, attention, and monotony. Rauterberg's adaptation is still called incongruity theory, but differs from Hunt's original proposals by shifting its focus toward improving understanding of the way humans interact with complex software systems. Rauterberg (1998) modified his application AMME to simulate user behaviour at various levels of mental complexity. With this process Rauterberg effectively created a user model to test his software on useability.

3.1.4 Modern uses of incongruity theory

Modern uses of incongruity theory involve attempts to improve player experience in games (van Lankveld et al., 2008), as well as attempts to analyse games (Halim et al., 2010). Recent publications usually focus on the differences between human skill and software complexity (most often in the form of game difficulty). They tend to ignore learning effects during software (or game) use. Nakatsu et al. (2005) builds upon the original incongruity work and extends it to a framework for entertainment in games. This framework consists of two dimensions being used to describe entertainment: the activity dimension and the integrated dimension.

Incongruity was used as inspiration for genetic algorithms that learned rules to enhance enjoyment in the CREATIVE COMMONS game (Baba and Handa, 2007). However, not much attention was given to empirically validating if human users actually experienced shifts in enjoyment due to differences in incongruity conditions.

Incongruity is often coupled directly to flow (Csikszentmihalyi, 1988), because both incongruity theory and flow theory revolve around emotions such as boredom, frustration, and pleasure. Cowley et al. (2008) have developed a user model that is aimed at detecting the amount of flow a user in a video game is experiencing. The model can provide both game-specific and game-independent predictions about player learning and entertainment. The model is empirically validated in the game of PACMAN [2] using 37 human participants (see Cowley et al., 2009). The algorithm used attempts to predict a user's future moves based on move data from previous game plays. The results are promising with the predictive validity of the model averaging out at roughly 70% chance of successfully estimating the next move. However, the model, as tested, does not seem to account for learning effects.

The experiment by van Aart et al. (2010) is an example of the modern use of incongruity. A mixed media experiment was set up in order to investigate if boredom and curiosity could be elicited by manipulating the scene in which a user was active. The goal of the experiment was to trigger boredom-related and curiosity-related behaviour. The researchers succeeded in eliciting the desired behaviours. The behaviour to be elicited in the curiosity condition would be that the subjects would follow a robot made to look like a rabbit. The game

[2] Developed and published by Namco in 1980.

scenario was based on the book *Alice in Wonderland*[3]. The validity of the experiment may be questioned by the choice of the scenario, since many adults know the story and would be stimulated to make choices that fit the original story. Admittedly, most people are affected by an effect called social desirability (Fisher, 1993).

Halim et al. (2010) investigated the entertainment value of six games by proposing entertainment metrics for all six of the games and then having agents play the games in order to see which of the games is most entertaining. The research rests on the assumption that it is possible to create an absolute measure of entertainment by using a metric. While this assumption might be correct, evidence of human perception and attitudes suggests that most human opinions are influenced by a large number of factors. Because of the many influences, the human opinion is unlikely to be sufficiently stable to be represented by the straightforward metrics used in the game. Obviously, the connections of the assumptions can be validated by testing the game metrics against human opinion surveys of the games used.

A recent approach to the balancing of incongruity, also called game difficulty balancing, was made by Aponte et al. (2011). The authors propose using synthetic players to measure the difficulty of games objectively. They argue that before any modification of difficulty can take place an objective measure of the actual difficulty needs to be obtained. Aponte et al. define video games as combinations or sequences of predefined challenges. They started by providing a description of difficulty in games, then they proceed to provide ways of testing the difficulty through the use of synthetic players.

Definition 3.5 (Difficulty balancing) Difficulty balancing is defined as the process of balancing a game to fit appropriately to the skill of a player.

3.2 The game: glove

For our experiments with incongruity we developed a game called Glove. It is an updated version of the classic game Gauntlet [4]. Glove contains a novel approach to keep incongruity at a desired level. For our first experiments, which are discussed in this chapter, we tested our approach, and investigated whether a large incongruity is less entertaining than a small incongruity.

In this section we describe the Glove game world (3.2.1). In Subsection 3.2.2 we describe the knight character. In Subsection 3.2.3 we describe the enemies in the game. In Subsection 3.2.4 we describe the balancing mechanism built into the game.

3.2.1 Game world

Glove (depicted in Figure 3.2) is a turn-based game, in which the player controls a knight. The knight is placed in a world that consists of cells. The world is 10 cells high, and 200 cells wide. Each cell is either passable (grass), or impassable (water or mountain). The knight occupies one cell. The world also contains enemies, each of which also occupies one cell.

[3]Written by Charles Lutwidge Dodgson under the pseudonym Lewis Carroll in 1865.
[4]Developed by *Tengen*, 1987

Figure 3.2: The game Glove. On the left side we see the knight. On the right from top to bottom, we see the dragon, the ninja, and the witch.

3.2.2 The knight

The knight starts at the leftmost end of the world (in one of the cells $(1, 1)$ to $(1, 10)$). His goal is to reach the rightmost end of the world. The game ends in victory for the knight (i.e., the human player) when the knight reaches the goal. It ends in defeat if the knight dies before reaching the goal. A knight dies when he has no health left. Health is measured in hitpoints, of which the knight has 100 at the start of the game. As soon as the number of hitpoints reaches zero, the knight dies. Each turn, the knight can take one of two actions: he can either move, or attack.

When moving, the knight leaves the cell that he occupied, and moves (either horizontally, vertically, or diagonally) over to any of the eight adjacent cells. Each move costs the knight 0.5 hitpoint. This means that if he moves steadily and unobstructed through the world, he has sufficient health to win the game, but has no health to spare.

When attacking, the knight executes an attack to one of the eight adjacent cells. He can either attack with his sword, or with a rock, which he may have picked up in the game world (by moving over it). The knight can carry at most one rock at a time. The difference between attacking with the sword and attacking with a rock, is that the rock attacks two cells, namely the cell which is attacked, and the one directly behind it (in the direction of movement). If an attacked cell contains an enemy, the enemy dies. The knight remains in its cell after attacking. Upon dying, the enemy leaves behind a health token, which the knight may pick up (by moving into the cell containing the token). Picking up a health token grants the knight 5 hitpoints (health can go up to a maximum of 100 hitpoints).

3.2.3 The enemies

There are three different types of enemies in the world, a number of which are spawned at regular intervals. Each time an enemy attacks the knight, the knight loses 5 hitpoints. The number of enemies in the game is determined by the balancing mechanism. The three enemy types are the following.

1. Dragon: The dragon approaches the knight using a shortest-path method. When the dragon is next to the knight, he may attack the cell that the knight is in. Arguably, the dragon is the easiest enemy to deal with for the player.

2. Ninja: The ninja has the same basic behaviour as the dragon, but has an additional ability: he can become invisible. He will use this ability when he is within a 5-cell distance of the knight, and will remain invisible for 10 turns, or until he attacks the knight. The ninja's behaviour is reasonably predictable, even when invisible, for players who possess a good mental model of the game.

3. Witch: The witch approaches the knight in the same way as the other two enemy types, but stops when she is within a distance of three cells of the knight. At that point, she will start to throw one fireball per turn in the direction of the knight. Fireballs move at a speed of one cell per turn. When there are few enemies on the screen, fireballs can usually be avoided easily. However, the knight must approach the witch to be able to attack her, at which time avoidance may be difficult. For most (but not all) players the witch is the hardest enemy to deal with.

Every enemy makes its move simultaneously with the player. This results in enemies and the player being able to hit each other at the same time. It also means that a player may hit a location where no enemy is yet, if he expects an enemy will move there when the action is executed. This is advanced player behaviour that usually takes a novice some practice to learn.

3.2.4 Difficulty and incongruity

The game of Glove has three difficulty settings: easy, balanced, and hard. Since players learn continuously during play, the difficulty needs to be kept constant by a set of algorithms. The algorithms that maintain the difficulty in the game are given in Appendix B. According to incongruity theory, the three possible difficulty settings should result in three different emotions: boredom, frustration, or pleasure. The experiment is designed to test the predictions of the incongruity theory.

Definition 3.6 (Easy) A state in the game of Glove is called easy when the player is able to finish the game with approximately full health.

Definition 3.7 (Balanced) A state in the game of Glove is called balanced when the player is able to finish the game with a small amount of health left or fail just before the end.

Definition 3.8 (Hard) A state in the game of Glove is called hard when the player is unable to finish the game and fails at approximately 50% completion.

3.3 Experimental setup

In this section we provide an experimental setup to measure the incongruity while a game is being played, and adapt and maintain the level automatically to the desired level of incongruity. This level should be one that the human player experiences as entertaining, regardless of skills and capabilities. Our goal is to investigate whether boredom and frustration are indeed associated with a decreased entertainment value and with increased incongruity.

3.3.1 Participants

In order to test the effect of our game balancing approach, we let a number of human test subjects play GLOVE. 24 subjects participated. Subjects received research credits for participating. The subjects' ages ranged from 16 to 31 years. All were Dutch native speakers. None of them had prior knowledge of the game. The subjects had a varying background, and varying experiences with computers and games. The exact subject background did not matter for this experiment, since the game balances itself automatically to the skills of the player. The balancing process is summarised above.

3.3.2 Procedure

Each human subject played the game four times. The first time was a training run, in which the player was able to experience the game controls. In this run, at each spawn point the same three enemies were spawned, namely one of each type. The player was allowed to interrupt this play whenever he wished. Once the training run had been finished the player was expected to start the actual experiment.

In the actual experiment, the subject played the game three times: once with an easy difficulty setting, once with a balanced setting, and once with a hard setting. The order in which the difficulty settings were presented to the subject was varied. The subject was not aware of the difficulty setting of the current game. After each game a digital questionnaire was presented to the subject.

3.3.3 Questionnaire

The questionnaire contained a total of 26 items. The items measure five emotional categories, namely boredom, frustration, pleasure, concentration, and curiosity. The boredom, frustration and pleasure items were chosen because they are emotions predicted by Rauterberg (1995). We added concentration and curiosity because we expected these emotions might also be influenced by changes in incongruity. Each item was administered using a Likert (1932) scale of seven points ranging from "does not apply to me at all" to "completely applies to me." The questionnaire can be found in Appendix A.

On the questionnaires, scores ranged from 0 to 6 on a Likert scale, which was assumed to be a continuous scale with an average of 3. For each subject, for each category, the average of the answers to the questions belonging to the category was calculated. Then, for each of the difficulty settings, the means of these averages over all test subjects were calculated. The means are presented in Figure 3.3 and Table 3.1.

3.3.4 Statistical techniques

For our statistical analysis of the results, we had to remove one subject from the pool because of an input error, leaving 23 subjects ($n = 23$). Normally, in order to compare means for variables, an ANOVA or t-test is sufficient. However, because we had three conditions (easy, balanced, hard) to predict the five variables (the emotional categories described in Subsection 3.3.3) and because we applied all three test conditions to each subject, a repeated measures MANOVA test was needed. Straightforwardly using multiple t-tests or ANOVA tests would have ignored possible interaction and repetition effects.

3.4 Results

Table 3.1: Descriptive statistics for each GLOVE difficulty.

Difficulty	M	SD	N
Boredom			
Easy	3.8636	.81184	23
Balanced	3.8182	.82441	23
Hard	3.6818	.79125	23
Frustration			
Easy	1.6136	.78197	23
Balanced	2.6364	1.32226	23
Hard	3.9773	1.23902	23
Pleasure			
Easy	3.1932	1.45556	23
Balanced	3.2386	1.35725	23
Hard	2.4886	1.29440	23
Concentration			
Easy	5.2091	1.10665	23
Balanced	5.3273	.90826	23
Hard	5.2000	1.02725	23
Curiosity			
Easy	2.2273	1.18771	23
Balanced	2.1818	1.22810	23
Hard	2.1364	1.11677	23

The repeated measures MANOVA multivariate test produced significant effects ($p < 0.01$). However, our analysis showed that there was no effect for concentration and curiosity so these items will be presented in the section results, but will not be discussed any further. After this, a post-hoc univariate analysis and contrast analysis were performed in order to examine the differences between the five measured variables and the differences of the difficulty on these variables. We found that the effect of order was not significant ($p > 0.05$). Because we feared that our results might be influenced by the previous experience which players have had with games, a subsequent analysis was performed to see whether there were significant effects of experience with computer games. This effect was also not significant ($p > 0.05$).

Next, we tested the effect of the difficulty setting on each of the five categories of the questionnaires. We found no significant results for the categories boredom, concentration, and curiosity ($p > 0.05$ for all of them). However, we *did* find significant effects for the categories frustration ($p < 0.01$) and pleasure ($p < 0.05$). The results are illustrated in Figure 3.3.

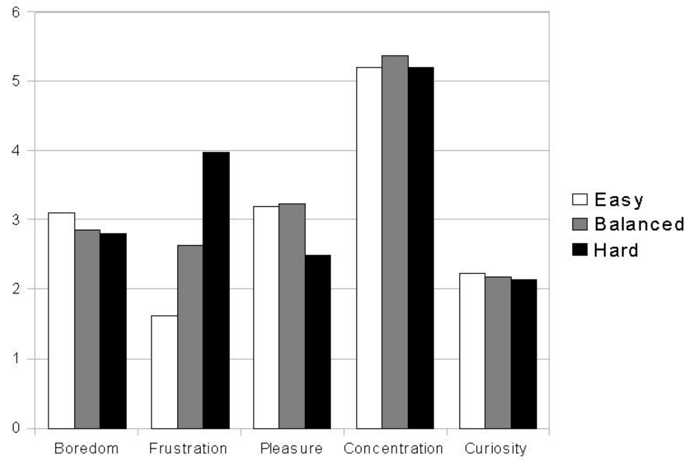

Figure 3.3: Incongruity questionnaire results.

Contrasts showed that for the category frustration, the differences between easy and balanced, and between balanced and hard difficulties were both significant ($p < 0.01$). Specifically, we found that GLOVE is significantly more frustrating for the balanced difficulty compared to the easy difficulty, and significantly more frustrating for the hard difficulty compared to the balanced difficulty.

For the category pleasure we found significant effects for the difference between balanced and hard difficulty ($p < 0.05$). In particular, we found that GLOVE provides significantly more pleasure for the balanced difficulty than for the hard difficulty. We did not find a significant effect for the difference between easy and balanced difficulty. The estimated marginal means for the category pleasure were 3.24 for easy difficulty, 3.25 for balanced difficulty, and 2.50 for hard difficulty. For the tables containing these results see Appendix D.

3.5 Discussion

In this investigation we pursued two goals: (1) to see whether boredom, frustration, and pleasure were related to incongruity, and (2) to investigate whether it is possible to manipulate what emotion a player will experience by automatically maintaining a high negative, high positive or low incongruity in our game.

Our tests show that our approach to game balancing, based on incongruity, can influence both the frustration level and the pleasure experienced in a game. The results show that a high positive incongruity is correlated to frustration, and that, at least for GLOVE, a balanced difficulty setting evokes more pleasure than a hard difficulty setting.

In summary, the results of our experiments show that incongruity theory is at least in part a consistent theory. The frustration effect follows the expectations of the incongruity theory, while boredom (which should be significantly higher for easy difficulty) does not follow the expectations. The pleasure effect follows the predictions of the incongruity theory in the balanced and hard difficulty settings.

It is likely that pleasure would also be as expected for easy difficulty, if easy difficulty was considered to be boring by the test subjects. Therefore it is interesting to examine why the easy difficulty setting was not found to be boring. We did not actually investigate this issue, but offer two possible explanations: (1) incongruity theory was originally applied to (relatively old) web interfaces (Rauterberg, 1994), and the increased visual and functional interactivity of our game, even in its simplicity, might cause a sufficiently high increase in complexity to be interesting in all modes of difficulty, (2) it is definitely possible that our easy difficulty setting is still sufficiently complex to create positive incongruity.

From the results we may tentatively conclude that our method of adaptive game balancing overcomes some of the problems of which commercial games suffer with their method of difficulty scaling, as our balanced difficulty setting manages to avoid the game becoming boring or frustrating.

In future work we may implement our adaptive game balancing approach in an actual commercial game, and test its effect on the pleasure value. Such an experiment could in particular demonstrate the applicability of our approach to commercial game developers, and may have an impact on how games are constructed in the near future. We note that the balanced condition may coincide with the balanced condition described in Iida et al. (2012), which also merits further investigation.

3.6 Chapter conclusions

In this chapter we examined (1) the relationship between game balancing and incongruity, and (2) how adaptive game balancing can be used to increase the entertainment value of a game. With respect to RQ1 *To what extent are games suitable for measuring incongruity?*, we may state that in our current experiment games are suitable for measuring incongruity. Incongruity theory states that emotions are felt when incongruity arises. We are able to find two of the five predicted emotional reactions in relation to the incongruity level (see the results section).

For our game GLOVE, we found that frustration increases with difficulty, while the pleasure value remains roughly the same for easy and balanced difficulty, but drops drastically for hard difficulty. From these results we may draw two conclusions: (1) incongruity theory is established as far as positive incongruity is concerned and (2) our approach to adaptive game balancing is suitable to maintain a game's pleasure value by keeping incongruity at a balanced value.

The pool of test subjects used for our experiments was relatively small, yet the results on which we base our conclusions are highly significant. However, we could not discover significant results for all the categories examined. Significant results for the remaining categories might be obtained with a higher number of test subjects.

In this chapter the player model used was purely a skill-based model, where skill is expressed as incongruity. Skill is relatively easy to measure as it is expressed in how well the player is able to deal with the challenges that the game offers. As far as pleasure is concerned, a balanced challenge should translate into high pleasure – which was partly established here. However, the pleasure value is also related to player personality, and personality is not so easy to measure in a game. Therefore, starting with the next chapter, we will investigate to what extent a player model can reflect player personality.

Chapter 4

Extraversion in games

In this chapter[1] we investigate RQ2: *To what extent can games be used to measure complex psychological traits such as extraversion?* We investigate whether games are suitable for testing psychology concepts related to behaviour that is more complex than the incongruity-related behaviour described in Chapter 3. We investigate the construct "extraversion". Extraversion is a concept from trait personality theory and it is known to emerge in a large number of situations. Extraversion was the first major trait to be found by personality researchers. To alleviate the drawbacks of the personality tests in use today (described in Subsection 2.5.5), our research aims to create an *automatic observational* test that is contained in a game.

In Section 4.1 we describe extraversion and we provide a short introduction to the field of psychological profiling and the reasons why we believe that a new way of testing would be a welcome addition to the currently available tests. In Section 4.2 we describe the game (letter to the king) used in this investigation. In Section 4.3 we describe the experimental setup used for investigating extraversion in games. In Section 4.4 we present our results. In Section 4.5 we discuss our findings. In Section 4.6 conclusions are formulated and answers to RQ2 are given.

4.1 Extraversion

The factor extraversion was first proposed by Jung, who described it as the inward or outward focus of libido. According to Jung (See Smith, 1977), people with low extraversion tend to turn their energy, focus, and orientation towards themselves, while people with high extraversion focus outside themselves. In contrast, Costa and McCrae (1992) describe people with high extraversion as sociable, meaning they prefer to be in the company of others and in social situations. They introduced the following six facets of extraversion.

[1] Based on: van Lankveld, G., Schreurs, S., Spronck, P. H. M., and van den Herik, H. J. (2011a). Extraversion in games. In van den Herik, H. J., Plaat, A., and Iida, H., editors, *CG'10 Proceedings of the 7th international conference on Computers and Games*, pages 263–275. I would like to recognize the publisher and to thank my colleagues for their permission to reproduce parts of the article in this thesis.

- *Activity:* Active, energetic people have high pace and powerful movement. They need to be busy and radiate a feeling of energy. They have a busy and hasty life.

- *Assertiveness:* Assertive people are dominant, self-confident, and controlling. They talk without hesitation and often lead groups.

- *Excitement-seeking:* Excitement seekers desire adventure, stimulation, and action. They like bright colours, noisy environments, and aculeated sensations.

- *Gregariousness:* Gregarious people prefer the company of others. They seek out others and like crowds and group activities.

- *Positive emotion:* People with positive emotion have fun, and feel happy and joyful. They laugh easily and are often cheerful and optimistic.

- *Warmth:* Warm people desire to form emotional bonds with others by showing warmth and affection. They are friendly and show that they genuinely like others.

Definition 4.1 (Extraversion) Extraversion is a personality trait representing the tendency to experience positive emotions, to crave activity and excitement, and to focus one's attention on others instead of on oneself.

The facets provide interesting information on their own but should always be considered in relation to the other facets and the factor as a whole (Costa and McCrae, 1995). Low scores on a facet do not indicate the opposite of the facet, just the absence of the tendencies of that facet. For instance, low positive emotion does not mean unhappiness, just an absence of positive emotion. In Subsection 4.1.1 we discuss automated player profiling. We relate extraversion to player modelling and profiling in Subsection 4.1.2.

4.1.1 Game observations

The benefits of an *automatic observational* test are that it has the potential to be accurate, truthfull, and inexpensive. An alternative potential benefit is that our game-based test can be an implicit test. In an implicit test, it is not immediately apparent to the test subject what is being measured. We are motivated by the fact that the function of the test is to measure personality, silently reducing the need for high human effort. In the recent past, an automatic observational test was considered to be virtually impossible (McCrae and Costa, 2003). With the current means we believeave it is possible: computations are fast, storage is large, and algorithms are accurate. An important requirement for the acceptance of automated observational tests is that they are validated.

Finally, we remark that for automatic player profiling several possibilities exist, such as using game environments and web environments. We chose to use a game environment for the following reason. Game environments provide the opportunity for players to engage in activities analogous to the real world, whereas web environments impose constraints on the human interface.

4.1.2 Player modelling versus player profiling

The definition for player modelling has already been provided in Chapter 1. Here, we also provide a definition for player profiling. An example of the use of player modelling techniques is to improve gameplay by adjusting difficulty or a storyline to the player's preferences.

Definition 4.2 (Player profiling) Player profiling is an automated approach to generating personality profiles such as the profiles generated by the NEO-PI-R personality questionnaire.

An example of a personality profile is the set of traits described in Subsection 2.5.3. In player profiling we look for correlations between the player's in-game behaviour and his scores on a personality test. This can be seen as a form of classification in which the classes consist of combinations of scores resulting from the five personality factors.

The major differences between player modelling and player profiling lie in the features modelled. Player modelling attempts to model the player's playing style, while player profiling attempts to model the player's personality. The models produced by player profiling are readily applicable in any situation where the five factor model can be used. Previously, Drachen et al. (2009) used unsupervised learning techniques to classify players into one of four types based on gameplay data. Canossa (2008) suggests designing flow experiences in games based on four subpersonality types that players may display while playing a game (killer, achiever, socializer and explorer).

4.2 The game: letter to the king

For this experiment we developed a game using the Neverwinter Nights environment. Neverwinter Nights is particularly suitable for this purpose, as it comes with a powerful, easy-to-use tool set called Aurora, that allows for the creation of large virtual worlds with social interaction and conversation. It also allows for the logging of player behaviour and player choices. In Subsection 4.2.1 we present the story of the game. In Subsection 4.2.2 we present the game's controls. In Subsection 4.2.3 we provide a list of the in-game elements which we use in order to gather data.

4.2.1 Story

We created a short story for the Neverwinter Nights module. Playing through the story takes about half an hour. The story starts with a little girl asking the player to deliver a message to the king. The road to the king is filled with obstacles and encounters. Examples are: a beggar, several guards, a cleric, and the townspeople. In the end, the player will meet the king, and the game ends upon delivery of the message. While the player works through this story he unknowingly provides behavioural data on 21 different in-game elements.

4.2.2 Controls

Neverwinter Nights is a top-down role-playing game. The player can see himself from an eagle-eye perspective. The player chooses a spot to move to by clicking somewhere on the ground. He can also interact with objects and game characters by clicking on them or he can start communication with them. Whenever the player's character moves the game's camera moves in order to keep the player's character in the center of the screen.

4.2.3 In-game elements

In this research we want to relate behavioural observations to written tests. We set out to produce in-game elements that allow us to observe behaviour automatically. When trying to create in-game elements, we established that directly converting items of the existing NEO-PI-R personality questionnaire into in-game elements is difficult. The NEO-PI-R asks introspective questions *about* behaviour. However, we need to construct in-game situations in which the player has the opportunity to display *actual* behaviour. As a source of inspiration to overcome this obstacle we studied the written test statement guidelines by Costa and McCrae (1992) and the extraversion experiments by Geen (1984). As a result, we defined our in-game elements to be based on NEO-PI-R statements as well as on real life situations that were expected to elicit extravert and introvert behaviour. Our items were designed in such a way that they give the players a broad range of possible behaviours and facilitate them in acting in a personal and natural way.

Definition 4.3 (In-game element) An in-game element is a variable in a video game that measures the frequency of occurrence for one specific behaviour performed by the player.

We divided the envisaged set of in-game elements into three categories: (1) choice and **A**ction, (2) implicit **B**ehaviour, and (3) **C**onversation. These categories served as guidelines for creating in-game elements for different types of behaviour. We attempted to create at least one in-game element for each combination of facet and category. Combining the three categories with the six facets means that at least 18 observation elements were needed, but we arrived at three more, for a total of 21. Details follow below.

- *Choice and Action* (A) encapsulates explicit and rational behaviour. The player faces a number of choices by in-game elements that range from choices which a high extraversion person would make to choices which a low extraversion person would make.
- *Implicit behaviour* (B) covers unconscious behaviour that is performed as an automatic preference. The in-game elements often involve (1) measuring the time a player takes to make a decision or (2) the distance that is travelled within a certain amount of time.
- *Conversational items* (C) represent conversational preferences. Differences in in-game elements can be found in context information, presentation, and style.

All in-game elements are sorted by facet of extraversion. As listed earlier (see Section 4.1), the facets are Activity (Act), Assertiveness (Ass), Excitement seeking (Exc), Gregariousness (Gre), Positive emotion (Pos), and Warmth (War). The items are coded as a combination of (1) the facet measured and (2) the category used. For example: GreB is an in-game element measuring gregariousness (Gre) by implicit behaviour (B). Below, we provide a complete list of the in-game elements we have created.

Activity (Act)

- ActB_1: The time it takes the player to complete the entire experiment.
  Active people are expected to finish the game faster.

- ActB_2: In the game, the player is forced to wait in a large, empty room for one minute.
  Active people are expected to walk around more than less active players (i.e., this means to cover more in-game distance during this period).

- ActC_1: The player is requested to wait.
  Active people are expected to respond less positively to this request.

- ActC_2: The player is asked to confirm his response on ActC_1.
  Active people are expected to stick to their choice.

Assertiveness (Ass)

- AssA_1 = In the courtyard the player is told that he needs a guard to escort him across. The player can respond with: - So be it, lead the way, I'll be right behind you.
  - Alright, I'm going to the castle, you can follow me if you want.
  Assertive people are expected to choose the second option.

- AssB_1 = The player has to talk to an NPC who is in conversation with another NPC. When the player addresses the NPC, the NPC tells him that he has to finish his conversation first. The conversation is looped however, and the time the player waits before breaking in on the conversations is measured in intervals of six seconds.
  Assertive people are expected to interrupt the conversation early after discovering the loop.

- AssC_1 = The player meets the king and has urgent information to tell him. When he addresses the king, he can choose three ways to do this:
  - Good day your highness, would you please listen to my story?
  - Greetings my king, I am sorry to have interrupted you like this, but there is something you need to know.
  - You need to listen to me, I have important information for you.
  Assertive people are expected to choose the last option.

- AssC_2 = After the player has given a beggar some money (or none), the beggar will ask for more. The player has three possibilities to answer, varying slightly depending on what the player gave the beggar:
  - Please leave me alone, I don't have anything (more) for you.
  - Sorry, that's all I can spare right now. / I'm sorry, but I don't have any money for you.
  - Look, (I already gave you something/I'm not giving you any money), so leave me alone.
  Assertive people are expected to choose the last option.

- AssC_3 = After AssC_2, the beggar is still not satisfied and gets aggressive and demands some money. This varies slightly depending on what the player said before, the possibilities to answer are all similar to:
  - Please, just leave me alone, I have to go somewhere.
  - You won't get my money.
  - I'm not giving you anything, try taking it if you dare.
  Assertive people are expected to choose the last option.

Excitement seeking (Exc)

- ExcA_1 = The player enters a room which he can alter to his liking. He can choose different coloured beams of light for in the room. He can choose:
  - No light
  - Blue, red, yellow or white light
  - Three different colours of light (red, yellow, blue)
  Excitement seeking people are expected to choose the last option.

- ExcA_2 = The player enters a room which he can alter to his liking. He can choose different kinds of music to play in the room. He can choose:
  - No music
  - Relaxing music (Stanley Myers - Cavatina)
  - Normal or pop music (DJ Disse and Batina Bager feat. Fred Astaire - Cheek to Cheek)
  - Upbeat music (The Prodigy - Colours)
  Excitement seeking people are expected to choose the last option.

- ExcB_1 = The player has to change clothes before meeting the king. He can choose three different outfits:
  - A black suit
  - A red suit
  - A multi-coloured suit
  Excitement seeking people are expected to choose the last option.

- ExcB_2 = After denying the beggar more money, the beggar will challenge the player physically. The player can then choose to (1) Fight or (2) Run Away.
  Excitement seeking people are expected to choose to fight.

- ExcC_1 = When the player brought the information to the king, the king thanks him and tells the player he is done and can wait in the next room. The player can respond in three ways, asking for more errands to do or being okay with waiting:
  - Alright, I'll be waiting in my room if you have anything else I can do.
  - Isn't there anything else I can do?
  - Let me do some more work, I've come all this way to warn you.
  Excitement seeking people are expected to choose the last option.

Gregariousness (Gre)

- GreA_1 = The player has to find some information about a girl for a guard. The player can go look in either (1) the library or (2) the bar.
Gregarious people are expected to look in the bar.

- GreA_2 = When the player reaches the courtyard, the guard there offers to accompany the player to the other side. The player can (1) decline or (2) accept this offer.
Gregarious people are expected to accept.

- GreB_1 = The player enters the bar and has to look for information. He has to find this by talking to the customers. These are spread into three differently sized groups. The group the player talks to first is considered his preference. There is (1) someone on his own, (2) two people talking, and (3) a big group.
Gregarious people are expected to talk to the big group first.

- GreC_1 = The player has to talk to a guard to ask him to open a door. The guard asks the player why he is in such a rush. The player can respond in three different ways, ranging from one where he wants to continue as soon as possible to one where he stays to chat a bit:
- I can't tell you, I have to hurry.
- I have some information that needs to reach the king as soon as possible.
- There is a girl who overheard some people talking about conspiring against the king, they said they were going to overthrow him.
Gregarious people are expected to choose the last option.

Positive emotion (Pos)

- PosA_1 = When the player is given the task to reach the king, the NPC asks the player the chances he will complete this. The player can answer in three ways:
- I will try, but it's not easy to get to see the king, so don't get your hopes up.
- I have no idea, I've never been in the castle before, but I'll do my best.
- It should be no problem, I am very resourceful. I'll be at the king in no-time!
Positive people are expected to choose the last option.

- PosA_2 = The player will pass a guard who asks if he can buy the players drink. The player can then ask three different amounts:
- Just take it for free, I'm not thirsty anyway.
- Sure, you can have it for 2 gold pieces.
- I'm thirsty myself, but you can have it for 5 gold pieces.
Positive people are expected to choose the last option.

- PosC_1 = After the player had the opportunity to alter the lighting and music in his room, the NPC outside asks him how he liked the room. The player has three answer possibilities:
- It's a bit sober, but it's good enough.
- It looks alright, and the music is quite good.

- Yes it's fantastic and the music is great.
Positive people are expected to choose the last option.

- PosC_2 = When an NPC in the bar asks the player how he is doing, he can respond in three different ways, in which the amount of interest in the NPC and the expression of the players own feeling is different:
  - I'm okay. Can I ask you a question?
  - I'm fine, thanks. How are you?
  - I'm feeling great! How about yourself?
  Positive people are expected to choose the last option.

Warmth (War)

- WarA_1 = A beggar comes up to the player in the courtyard and asks for some money. The player can choose three different things to do:
  - No sorry.
  - Here, some small change for you.
  - Of course, here's a gold coin for you.
  Warm people are expected to choose the last option.

- WarB_1 = In the bar, the player has to talk to at least five NPCs before they get the information they were looking for. After getting that information, the player can continue to the next room. There are a few more NPCs in the room however, and for every NPC the player talks to after the information is known, a point is added. Warm people are expected to talk to NPCs.

- WarC_1 = In the bar an NPC will talk to the player and asks him how he is doing. The player can respond in three different ways, in which a different amount of attention is given to the NPC:
  - I'm fine, can I ask you a question?
  - I'm fine. How are you?
  - I'm fine, thanks. Can I buy you a drink?
  Warm people are expected to choose the last option.

Added a posteriori

During our analysis we noticed that some players skipped conversation with a beggar character completely. In order to represent this behaviour we added the "skipped" in-game element.

4.3 Experimental setup

Our claim is that a player profile can be constructed by automatically observing the player's behaviour in a game. To test our claim we investigated the correlation between a person's game behaviour and his scores on a personality questionnaire. The experiment consisted of three phases: (1) subjects play a game, (2) subjects complete a personality questionnaire, and (3) subjects complete an additional questionnaire containing topics of possible relevance to the experiment. In Subsection 4.3.1 we discuss the extraversion experiment. In Subsection 4.3.2 we discuss the participant in our experiment. In Subsection 4.3.3 we discuss the experimental procedure we followed. In Subsection 4.3.4 we discuss the statistical techniques we used to analyse our data.

4.3.1 The extraversion experiment

Participants to the experiment entered in one of two conditions, either (1) subjects were invited to rate the statements of the NEO-PI-R on extraversion and then to play the game followed by a general information questionnaire, or (2) the participants played the game followed by the NEO-PI-R on extraversion and finally the general information questionnaire. The topics of the general information questionnaire included age, gender, and experience with computers and games.

4.3.2 Participants

A pool of 39 participants, containing 20 males and 19 females, was tested. Ages ranged from 18 to 43 with a mean age of 24. Most participants were either students or former students. Participants received research credits for participating. All subject data was processed anonymously.

4.3.3 Procedure

Answering the personality questionnaire took 10 minutes. Playing the game took between 30 and 40 minutes. Answering the additional questionnaire took 10 minutes, too. Each subject tested for a maximum of one hour. For playing an instruction booklet was provided, asking participants to respond if possible as they would do in real life. Instructions on playing the game were included in the booklet which can be found in Appendix E.

4.3.4 Statistical techniques

The results (Section 4.4) were analysed by SPSS using a multiple linear regression analysis. The NEO-PI-R returns results on a 1 to 9 scale. Correlations were calculated using extraversion; the facet scores were used as dependent variables and the 21 in-game elements as independent variables. Furthermore, regression analysis was conducted to inspect the relationships between the control variables and the extraversion scores.

4.4 Results

Our claim is that the in-game elements have a correlation with the facet and extraversion scores of the NEO-PI-R. Therefore, the questionnaire answers should function as predictors for extraversion and its facets. Our experiments were meant to investigate to what extent this happened.

The results of the experiment are summarised in Table 4.1. On the horizontal axis, the table contains the factor extraversion and its facets. On the vertical axis the table contains 12 of the 21 in-game elements, namely those that showed some correlation with one or more of the facets or extraversion itself. We denoted the effect size by r and the significance by p.

Because of the large variations commonly present in human behaviour and the large number of factors influencing this behaviour (personality, intelligence and learned associations) psychologists consider the following correlations to be indicative for effect sizes in a relationship between personality and the participants' game behaviour (Cohen, 1988, 1992).

- *small effect* $r = .10$ (1% of variance explained)
- *medium effect* $r = .30$ (9% of variance explained)
- *large effect* $r = .50$ (25% of variance explained)

In the table there is a distinction between positive correlations and negative correlations. A negative correlation indicates an inverse relationship between a factor and its facets and an in-game element. If the in-game element increases in value its related facet decreases.

Table 4.1 contains the correlations between (1) in-game elements and (2) the NEO-PI-R scores. It should be noted that the in-game element named "skipped" is added to the table. This was done because some of the subjects broke off the conversation with the beggar (a character in the game). After a closer investigation it became apparent that the players concerned had skipped the beggar accidentally. Skipping the beggar was significantly related to a low control skill in the game ($p < .05$).

Table 4.1: Correlations between NEO-PI-R scores and game items.

	Extraversion	Act	Ass	Exc	Gre	Pos	War
ActB_1			.327*				.279*
ActB_2		-.279*					
ActC_1	.321*	.339*	.303*			.269*	
ActC_2	.271*		.351*		.451**	.293*	
AssA_1						.302*	
AssB_1				.353*			
ExcB_1	-.318*			-.325*	-.349*	-.302*	
GreA_2	-.321*			-.605**			
GreB_1				.432**			
PosA_1	.307*		.294*				
WarC_1						.278*	
Skip				-.277*			

** $p < 0.01$, * $p < 0.05$

Control questions

The control questions can be found in Table 4.2. Additionally, Table 4.3 contains (1) the correlations beween the control items and (2) extraversion and the game items. Its columns stand for: sex, age, education, experience with computers, experience with games, English language skill, ease of the controls, and clarity of the in-game missions.

Table 4.2: Control questions.

Question	Possible answers
Age	select an age
Sex	male or female
Education level	select an education level
Describe your level of computer experience	1 (very low) to 5 (very high)
Describe your level of gaming experience	1 (very low) to 5 (very high)
Describe your level of English language skill	1 (very low) to 5 (very high)
Please rate the ease of use of the game controls	1 (very difficult) to 5 (being very easy)
Please rate the clarity of game missions	1 (very unclear) to 5 (very clear)

Table 4.3: Correlations between control questions and game items.

	Sex	Age	Edu	ExpC	ExpG	Eng	Eas	Cla
Extraversion					-.417*			
ActC_2		.344*		-.364*				
ExcA_1	-.462*			.518**	.469**			.518**
ExcB_2					.347*		.356*	
ExcC_1			.364*					
GreA_1							-.420*	
GreA_2							.394*	.393*
GreB_1							.353*	
GreC_1						.376*		
PosA_1								.360
PosC_2		.355*			.360*			
WarC_1	-.376*							

** $p < 0.01$, * $p < 0.05$

Table 4.3 shows that a large number of effects were found in the control questions. Elements such as age, sex, experience with computers and games, and skill of interacting with the game seem to be correlated with many of our test items and even with extraversion itself. For instance, it seems to be the case that experience with games is indicative for lower extraversion scores, which underlines the stereotype of the "introverted gaming nerd". This means that values for test items, facets, and extraversion might be derived not only from observing a player's behaviour in the game, but also from his handling and understanding of, and attitude towards the game. It also means that, in future work, we might need to correct the results derived on test items for the meta-information from the control questions.

4.5 Discussion

In this section we interpret the results and discuss them. As stated at the start of the chapter, the goal of the present research was to model a subject's personality automatically. We based this model on the player's in-game behaviour, i.e., his actions and choices in a game. The drawback of using a game is that players can act unlike their 'real-life personality' and more like the role of the character that they wish to play. However, we assume that, even if players are acting according to their character's role, there will still be a substantial number of characteristic behavioural patterns that result from their personality. We discuss the results for extraversion in Subsection 4.5.1 and the effect of our significance level in Subsection 4.5.2.

4.5.1 Extraversion

The NEO-PI-R results show that our test subjects scored above average on extraversion. The scores range from 1 to 9 with 4 as the lowest measured score in the group of participants. Table 4.1 shows the significant correlations between five of the in-game elements and extraversion. Three of the correlations are positive and two are negative. All correlations are significant on a level of $p < 0.05$ or lower. Items ActC_1 and ActC_2 were conversation elements involving the willingness to wait, and item GreA_2 represents the choice of having a guard accompany you across a courtyard or not. Item ExcB_1 is the choice of colourful clothing which was scored from low being black to high being quite colourful. PosA_1 is a conversation element displaying the amount of optimism when asked whether the player believes that the game mission will be a success. Three of the five in-game elements showing correlation are conversation elements, one is an implicit and one is an explicit choice. None of the other 21 in-game elements showed any correlation sufficiently high to be significant for extraversion, but the 12 elements in Table 4.1 showed correlation with the facets.

12 of the 21 in-game elements demonstrated correlation with extraversion or with its facets. We have found a total of 24 significant correlations, three of which reach the $p < 0.01$ level of significance. Our expectation was that each of the in-game elements would correlate with a given facet. However, we found that while each facet has strong correlations to in-game elements, only the facets activity and excitement seeking correlate to *their own* in-game elements.

Seven of the 24 correlations that were found had a negative value. This indicates that they produced an effect opposite to what was expected based on the personality literature the in-game elements were based on. Currently, we have no explanation for these effects.

4.5.2 Significance level and control variables

There are three important caveats that we need to clarify regarding the interpretation of correlation and experimental control. Each statistic runs the risk of misinterpretation. Below, we attempt to clarify these risks.

Correlation

In significance analysis, there is a risk of accepting a correlation as significant while it is not significant. A significance level of $p < 0.05$ means that on average 1 in 20 of the investigated correlations incorrectly shows up as significant. At $p < 0.01$ we have incorrectly assume significance of an average of 1 in 100 correlations.

We note that in this research we have 21 in-game elements and that we analyse the correlations of these elements with seven constructs (extraversion and six facets). The resulting analysis counts $21 * 7 = 147$ correlation analyses. For 147 correlations (the number in our investigation), an average of 7.35 is therefore incorrectly classified as significant. There are three correlations in our results that reach the $p < 0.01$ level of significance.

(Field, 2009) proposes five solutions for the problem inherent in significance levels (1) increase the number of participants, (2) use a stricter correlation test, (3) use a stricter significance level, (4) test less variables, and (5) use crossvalidation with new test participants. We did not have the opportunity to increase our number of participants and a stricter test and stricter levels result in only three correlations being left. Less variables was no a posteriori option so we refer to Chapter 5 for a setup that attempts crossvalidation with new participants and a different approach as well.

Experimental control

We note that the control questions are not used as statistical controls in the correlation tests. With a population of 39 participants this means that at most, three control questions could have been used. Given the correlations found here, in future work the use of control variables should be considered when using larger samples.

4.6 Chapter conclusions

In this chapter we investigated RQ2: *To what extent can games be used to measure complex psychological traits such as extraversion?* In order to investigate this research question we designed a test that measured extraversion and its facets in a game. We created a set of 21 in-game elements for the game Neverwinter Nights. The in-game elements were based on the questions of the NEO-PI-R. They were divided into three categories: choices and actions, implicit behaviour, and conversation. Investigating the question of correlation between in-game behaviour and personality scores on the NEO-PI-R, a test was administered to a pool of 39 participants and yielded outcomes for the 21 in-game elements. The outcomes were analysed for correlations using regression analysis. From the results we may conclude that it is possible to measure extraversion and its facets, using behaviour in a virtual world.

Table 4.1 shows that five of our in-game elements had a significant direct correlation to extraversion scores on the NEO-PI-R. Seven in-game elements had a correlation to one or more of the facets of extraversion rather than to extraversion directly.

In the following chapters, we will expand our research to include the other four factors of personality, to compare the predictiveness of player profiling to written personality tests.

Chapter 5

Data-driven personality in games

In this chapter[1] we investigate RQ3: *To what extent can we use games in order to create a full personality profile automatically?* In our experiments we use a module for the game NEVERWINTER NIGHTS (NWN) as a personality assessment tool. We examine whether individual differences in video game behaviour are related to differences in personality. We do so by correlating recorded game behaviour to scores on the NEO-PI-R personality questionnaire. A scenario is used that is similar to the analysis of those correlations found in the literature on today's commercial computer games.

In Section 5.1 we provide information on personality in games. In Section 5.2 we describe the game and module used in our experiment. In Section 5.3 we provide our experimental setup. In Section 5.4 we present our results, which are discussed in Section 5.5. Section 5.6 provides conclusions.

5.1 Measuring personality

According to Costa and McCrae (1995) personality (outside of games) is defined as the stable pattern of variation in individual acting, thinking, and experiencing. Personality arises from interactions between (1) the situation in which the individual is placed and (2) the processes that take place in the individual (Back and Egloff, 2009). Personality theory implicitly assumes personality results from interactions; personality scores are a result of measurements across situations and can therefore be generalised (ten Berge and De Raad, 2002).

In this thesis we focus on the five factor model of personality (FFM). A description of this model is found in Subsection 2.5.3. Commonly, personality is measured using three

[1]Based on: van Lankveld, G., Spronck, P. H. M., van den Herik, H. J., and Arntz, A. R. (2011b). Games as personality profiling tools. In *2011 IEEE Conference on Computational Intelligence and Games (CIG)*, pages 197–202. IEEE. I would like to recognize the publisher and to thank my colleagues for their permission to reproduce parts of the article in this thesis.

types of measures: questionnaires, interviews, and observations. These three measure types are described in more detail in Subsection 2.5.5. In this chapter we add a new measure to this list, automated personality profiling.

Definition 5.1 (Automated personality profiling) Automated personality profiling is defined as the automatic collection of gameplay data for the generation of a personality profile.

Automated personality profiling

For automated personality profiling, we created a module for the game Neverwinter Nights[2]. We chose to design a game scenario (henceforth module) in which the player experiences many of the situations commonly found in role-playing games. Our setup was meant for participants to experience a wide range of game situations in a time frame of 60 minutes. We expected 60 minutes to be a sufficient amount of time for gathering a representative sample of player behaviour.

5.2 The game: the poisoned lake

Neverwinter Nights is a top-down role-playing game in a medieval fantasy setting. The game is accompanied by a toolset called Aurora, which enables the design of modules through the placement of area tiles, characters, and objects. Our module contains the story, characters, and the relevant locations in the game. In order to show the experimental setting accurately, our description of the game includes (1) the controls used, (2) the world in which the story takes place, and (3) the story of the game.

5.2.1 Controls

Our aim was to keep the game controls as straightforward as possible in order to minimize the learning curve involved to master gameplay. To reach this goal we only use mouse movement and the left mouse button to control the game.

The interaction between a participant and the game is by mouse control. The player can (1) move by clicking in an area, (2) interact with objects by clicking on them, and (3) start conversations by clicking on a non-player character (NPC). Conversations are in the form of menus; the player chooses a response on an NPC statement from a list of possible answers (lists contain 1 to 5 items).

5.2.2 The game world

The world is made up of 16 areas. There are five outside areas and eleven inside areas. The outside areas are displayed in Figure 5.1. The inside areas consist of: (1) a dream (training area), (2) the player's house, (3) the top floor of his house, (4-6) three houses, (7) a shop, (8) an inn, (9) the top floor of the inn, (10) the tower inside, and (11) a cave. Except for the training area, the player can freely move between all areas once he has opened the way to them. The player cannot return to the training area because the training area is a dream

[2]Developed and published by *BioWare* in 2002

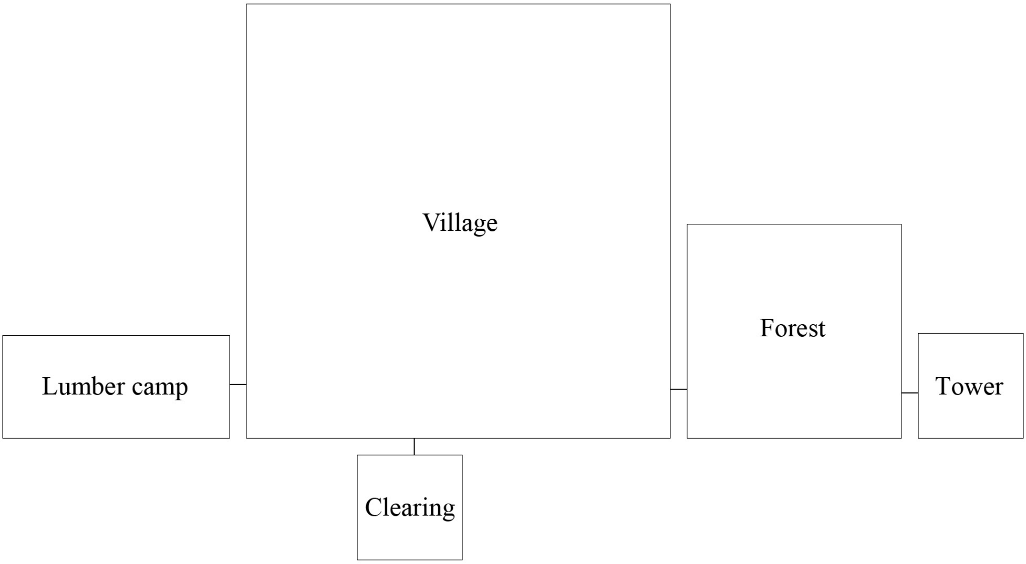

Figure 5.1: The five outside areas of the game world.

and once the player wakes up from this dream there is no option to return to sleep. There is only one way closed to the player. This is the way from the village to the forest and it is closed for story purposes. The participant can only overcome this obstacle after having found the dead shopkeeper (this obstacle is further explained in the story description). If the participant has not found the poisoned shopkeeper he has no reason to be in the forest and in the tower (after the forest) in order talk to a character there.

Table 5.1: A list of the characters in the game.

Character Name	Area	Story
MrRed	Dream	Main
MrBlue	Dream	Main
Siline	Lumber camp	Side
Dara	Village	Main
Old man	Village	Main
Burrick	Village	Side
Evana	Village	Side
Myztor	Forest	Main
Moricho	Cave	Main

5.2.3 The story

The game's story consists of three parts: (1) a training sequence, (2) the main story, and (3) two smaller side stories that are unrelated to the main story. Table 5.1 contains a list of the characters in the game that are important to the main and side stories. The table contains nine characters, there are a total of eighteen characters with which conversation is possible. The characters that are not directly involved in the main or the side stories are not included in the table. The table also shows whether the characters are involved in the main or side story. The game starts in a training area in which the participants learn how to perform the various actions that are possible in the game. Participants also learn how to use the map and their inventory. Moreover, in the training the use of the logbook is explained. The logbook records the events of the main story line and can be consulted in case the participant cannot discover how to continue. After the training, the participant starts the main story. The main story involves a multi-step mission that leads the participant through various situations commonly found in commercial video games. A short summary of the main storyline is as follows.

- Go to the village for an errand.
- Discover a poisoned shopkeeper.
- Go to a sage for advice.
- Go to the cave to stop the cause of the poisoning.

There are two side stories. The completion of these stories is not required for the completion of the main story. The side stories are only encountered if the participants take the time to talk to the NPCs that start the stories, Siline and Evana. In the first side story the participant has to go talk to an NPC to ask him to stop bothering Siline. The second side story starts when the poisoning is discovered. In this story, the participant has to tell a child to go home to its mother in order to reduce the risk of her being poisoned. Following these side stories will lead to a delay in completing the main story. Such a delay is one of the indicators for playing style, preference, or behaviour.

5.3 Experimental setup

In our experiment we investigated whether a correlation exists between personality scores and game behaviour. In order to perform our experiment we applied two measurements: (1) participants took the NEO-PI-R personality test, and (2) the same participants played the game and we recorded their behaviour. In Subsection 5.3.1 we present general information on the participants and the setup of our investigation. In Subsection 5.3.2 we give a description of the variables we have constructed in the game in order to measure behaviour in the game.

5.3.1 Participants and time frame

In total we had 80 Dutch speaking participants. Participants received research credits for participating. For one player, playing the game lasted 60 minutes at maximum (45 minutes average). If the player did not finish within 60 minutes the game automatically stopped. Completion of the full five factor personality questionnaire took a maximum of 60 minutes (average 45 minutes). The resulting complete experiment took a maximum of 120 minutes with an average of 90 minutes. The participants were informed that all data would be collected anonymously. The order in which the questionnaire and the game were given was reversed for half of the participants. This was done to counterbalance any effects that playing the game might have on responses on the personality questionnaire and vice versa.

5.3.2 Game variables

In this experiment we collected data of the participant's game behaviour. We describe two categories of variables, pooled and unpooled (see below). The number of pooled variables was 43 and the number of unpooled variables was 217, making a total of 260 variables for the entire game. All variables in the game are natural numbers with an unlimited range. Below we first explain the rationale behind pooled variables, which are characterised by the fact that they all belong to the same pool. Thereafter we explain the unpooled variables, which are characterised by the fact that they do not belong to a pool. The following list presents an overview of the final variable set we used. The variables are divided into five groups in order to clarify the results presented in the next section. The first four groups contain the pooled variables while the final group contains the unpooled variables. A list of all the variables and the number of variables per group can be found in Appendix F.

- *Group 1* contains four pooled game variables monitoring conversation and movement variables for the entire game.
- *Group 2* contains sixteen pooled move variables per area. Each pooled move variable contains the total movement for a specific area.
- *Group 3* contains seven pooled conversation variables per area. Each pooled conversation variable retains the total number of conversation choices for a specific area. There are seven areas in the game where no conversation can take place.
- *Group 4* contains sixteen pooled conversation variables (one per NPC). Each pooled conversation variable contains the total number of conversation choices for a specific NPC.
- *Group 5* contains the 217 unpooled variables.

Pooled variables

Pooling behaviour can be used as a way of improving the predictiveness of a behaviour. In some situations, psychologists find pooled variables more indicative than unpooled variables because pooled variables are considered to be expert knowledge, knowing which variables to pool requires the insight of an expert. In order to pool a variable, game variables that are influenced by the same psychological process should be selected. The pooling can be done by adding the values of the variables into a new "pooled" variable. An alternative to addition is to transform the variable values into z-scores and to take the mean of these z-scores. In adding values the relative influence of the original variables is maintained in the pooled variable. In creating z-scores each variable receives an equal amount of influence in the pooled variable. Choose the method which is most in line with the goals of the research.

Definition 5.2 (Z-scores) Z-scores are also known as standard scores. Z-scores are calculated by transforming all the measured experimental scores from their original distribution to fit a normal (or Gaussian) distribution with a mean of zero and a standard deviation of one.

Definition 5.3 (Pooled variable) A pooled variable is defined as a variable with a value that is comprised of the values of several other variables.

An important point of attention when pooling behaviours is that only behaviours should be pooled that represent the same construct. An example is talkativeness in games. In order to represent talkativeness in games we could pool each conversation into a variable called "total conversation time". Adding the values for movement to this variable would make no sense.

When we have created our "total conversation time" variable, it may be influenced by other traits besides talkativeness. For instance, if the game is an investigation game, a factor like curiosity may influence the number of questions asked. The pooled variable may then be representing not only talkativeness but also curiosity. Low talkative but highly curious participants may produce equal scores compared to highly talkative but low curious participants. The variable's validity may be tested by controlling for curiosity and other confounding factors or by testing if the low-level variables truly measure the concept talkativeness by testing for inter-correlatedness or correlation with an alternative talkativeness measure.

Pooled variables in our research

To comply with the experimental considerations we created pooled variables that combine the counts of all unique variables per area, per NPC, and for the entire game. These pooled variables could, for example, be used to examine the overall tendency of a player to move around or to engage in conversation. Individual unpooled variables might miss such tendencies. We created pooled variables by summing the values of several unpooled variables.

Unpooled variables

In order to gather raw behavioural data, 217 variables were created. They were split into three types of behaviour, (1) movement, (2) conversation, and (3) miscellaneous (e.g., interaction with objects). There were 92 variables that recorded movement behaviour, 120 variables that recorded conversation behaviour, and five variables that recorded miscellaneous behaviours. Each variable recorded the total number of times its monitored behaviour was performed. Conversation variables recorded choices made in conversations. Each time one of the conversation choices was made, the value of its respective variable increased by one.

Definition 5.4 (Unpooled variable) An unpooled variable is defined as a variable with a value based on directly measured behaviour.

Figure 5.2 contains an example of a conversation. For a proper understanding we have made a translation from Dutch to English. The conversation occurs when the participant first encounters the NPC Siline. Siline: *Hi Moris how are you?* Possible player responses are: (1) *I'm fine*, (2) *What are you doing here?* (3) *Who are you?* and (4) *I had the most bizarre dream.* Each unique choice option is recorded by a unique variable in our dataset. Since it is possible to revisit conversations and then make the same or a different choice each discrete choice is a frequency measurement ranging from 0 to a potentially infinite size.

Figure 5.2: A screenshot of a typical conversation.

The movement variables similarly record the total number of movement behaviours for each variable. The value of a movement variable was raised every time the participant entered the location in the game monitored by the variable. The movement variables were placed along the doors between the areas, halfway across areas, and around special objects such as trees and gardens. The variables were triggered whenever the player's character moves onto the trigger's location. In Figure 5.3 we can see three triggers. These triggers are represented by the blue boxes on the floor. These boxes are only visible in the editor, not in the game while it is being played. The first trigger is in the entrance of the door, recording entering or leaving movement, the second trigger has an angle and is slightly further inside the room which might be triggered if the player chooses to explore the room, and the third trigger is at the far side of the room and might be triggered if the player chooses to explore the room to its maximum extent.

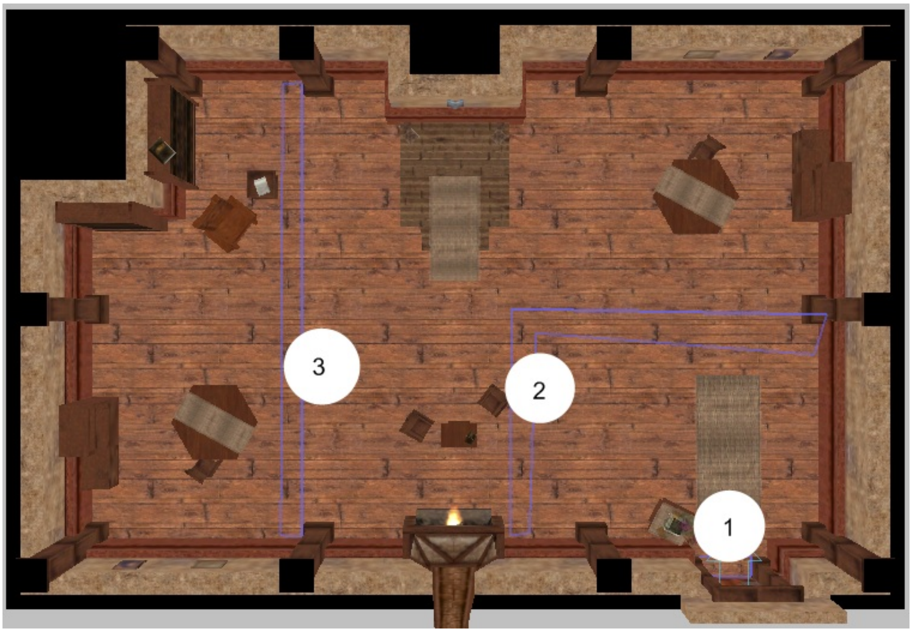

Figure 5.3: The interior of a house containing three movement triggers.

5.4 Results

We start with an overview of the data collected in this experiment. Table 5.2 contains the descriptive statistics of the OCEAN traits. Appendix F contains the descriptives of all the in-game variables (pooled and unpooled). The NEO-PI-R generates a stanine (value from 1 to 9) score for each trait. The table shows that the participants' trait scores ranging from 1 to 9 can be found for openness, conscientiousness, extraversion, and neuroticism. For agreeableness we find a range from 1 to 8. From these ranges, we may conclude that our participants lacked the most extreme cases of agreeableness. The averages of our participants' trait scores are quite representative for a normal population. In order to receive the full dataset please contact the author (they are available on request).

Table 5.2: Descriptive statistics.

Trait	N	Minimum	Maximum	M	SD
Openness	80	1	9	5.83	1.659
Conscientiousness	80	1	9	4.28	2.006
Extraversion	80	1	9	6.12	2.015
Agreeableness	80	1	8	3.66	1.835
Neuroticism	80	1	9	5.25	1.899

Linear regression analysis

We have first performed a linear regression analysis in order to see whether it is possible to fit a linear model to predict the OCEAN personality traits based on the pooled and unpooled game variables. We used stepwise linear regression. The results can be found in Table 5.3. Column 1 contains the OCEAN personality traits, column 2 contains the effect size (R^2) and column 3 contains the number of game variables used in the linear model. The effect size represents the "goodness of fit" of the model. "Goodness of fit" denotes the amount of variance explained by the model. Effect sizes range from 1 to 0. A value of 1 means that 100% of the variance is explained by the model, a value of 0.5 means that 50% of the variance is explained by the model, etcetera. In this linear regression analysis both the pooled and the unpooled variables have been included. Table 5.3 shows that we can most accurately predict variance for openness and agreeableness (around 75%) followed by conscientiousness and neuroticism (around 55%) with extraversion trailing (at around 35%).

Table 5.3: Stepwise linear regression between traits and game variables.

Trait	R^2	Variables in model
Openness	.768	17
Conscientiousness	.559	10
Extraversion	.351	6
Agreeableness	.724	15
Neuroticism	.568	9

Correlational analysis

For a more detailed investigation, all variables (pooled and unpooled) were analysed using a correlation analysis. Although the linear regression analysis calculations have already predicted whether correlations for the correlation variables are positive or negative we can still look at the two-sided correlations in order to have a higher degree of confidence in the results. We investigated the correlations between the game variables and scores on the five traits measured by the NEO-PI-R personality questionnaire. The analysis shows significant results for all traits of the FFM. Correlations with $p < 0.05$ are considered to be significant. Below, we describe the results by personality trait.

Table 5.4 contains the total number of positive and negative significant correlations per group. Empty cells indicate a lack of significant correlation. Table 5.5 contains the lowest and highest of positive and negative correlations per group. For interpretation of correlation effect sizes see Section 4.4. For the sake of clarity, non-significant correlations are not included in either table.

In Table 5.4 we observe that significant correlations are found between all five personality traits and game variables in all groups. We observe effects of openness in all five groups. 25 of the 27 correlations for openness are negative. Conscientiousness shows three positive effects and three negative effects (found in group 5).

For groups 1 and 3, openness is the only trait that has influence. In group 2 we see effects of openness as well as extraversion. In group 4 we see an effect of both openness as

Table 5.4: Total number of positive (pos) and negative (neg) correlations per group.

Trait	Group 1		Group 2		Group 3		Group 4		Group 5		Totals	
	Pos	Neg	Pos	Neg	Pos	Neg	Pos	Neg	Pos	Neg	Pos	Neg
O		3		2		1		1	2	18	2	25
C									3	3	3	3
E			1						2	2	3	2
A									8	5	8	5
N					1				7	3	8	3
Totals	0	3	1	2	1	1	0	1	22	31	24	38

The letters OCEAN stand for their respective traits.

well as neuroticism. In group 5, effects of all traits can be found.

We observe a total number of 24 positive correlations and 38 negative correlations. Some variables had a correlation to more than one trait. In total 57 unique variables had one or more correlations. This is 57/260 (or 22%) of all the variables. Each of the five personality traits is correlated to one or more game variables. From these results we may conclude that significant effects of all five personality traits are present in conversation and movement behaviour in the game.

In Appendix H we present the full list of correlations that produced significant correlations.

5.5 Discussion

In the discussion below we focus on six topics. We discuss: (1) the correlations per trait, (2) the effects of openness, (3) we discuss the different ways of interpreting the correlations that have been found, (4) the capability of our game to model extraversion, (5) interpreting regression results, and (6) effect sizes and significance.

5.5.1 Correlations per trait

We investigated the relationships between personality and video game behaviour. All five traits of the FFM, as measured by the NEO-PI-R have been found to correlate significantly with game behaviour. Below, we discuss our results per group of variables. Here, the word “significantly” refers to the psychological norm of significance (i.e., the found results occur by chance only one in twenty times or less). As stated in the previous sections, the results in the tables are all statistically significant. We found the following results per trait and per group. All the results in this subsection are based on examination of the Tables 5.4 and 5.5 and on Appendix H. In this subsection we also refer to a number of NPCs that can be found in Table 5.1.

Table 5.5: Highest and the lowest correlations for each group.

Correlation	Group 1		Group 2		Group 3		Group 4		Group 5	
	Pos	Neg	Pos	Neg	Pos	Neg	Pos	Neg	Pos	Neg
					Openness					
Lowest		-.234		-.229		-.310		-.312	.274	-.220
Highest		-.255		-.239		-.310		-.312	.300	-.377
					Conscientiousness					
Lowest									.223	-.230
Highest									.235	-.318
					Extraversion					
Lowest			.226						.229	-.229
Highest			.226						.254	-.240
					Agreeableness					
Lowest									.222	-.234
Highest									.384	-.304
					Neuroticism					
Lowest					.229				.226	-.233
Highest					.229				.265	-.310

Openness

For group 1 we see that three variables correlated negatively with openness (between -0.234 and -0.255). The negatively correlated variables record the frequency of (1) conversation starts, (2) conversation endings, and (3) the total number of movement-variable triggers in the game.

In group 2 we see that two variables correlated negatively with the openness trait (-0.229 and -0.239). The negative correlations show that a higher openness score relates to less movement in the village and lumbercamp areas.

When we look at group 3 we see that one variable correlated negatively with openness (-0.312). In the village area the players with high openness converse less.

Group 4 shows that one variable correlated negatively with openness (-0.310). This variable represents the total amount of conversation with a character in the village named Marto. The main topic of conversation with Marto is getting directions to locations in town. The Marto character is not in Table 5.1 because he is not relevant for either the main story or the side story.

In group 5, two variables correlated positively with openness (0.274 and 0.300) and 18 variables correlated negatively with openness (between -0.220 to -0.377). There does not seem to be a clear pattern to the correlations.

Large trends can be seen in the variables in the openness personality trait. From these results we may conclude that players with higher openness move around less in the game and make less conversation in the game. In summary, it seems that high-openness participants tend to focus mainly on quickly moving through the main story of the game and spending as little time as possible in conversation.

Conscientiousness

Group 5 shows that conscientiousness correlates positively to three variables (between 0.223 and 0.235) and negatively to three variables (between −0.230 and −0.318). The variables relate mainly to the completion of side quests.

Overall, the scores on conscientiousness do not seem to follow a coherent pattern.

Extraversion

In group 2, one variable correlated positively (0.226). This variable measures the amount of movement in the dream training area. The dream is the only area with surrealistic lighting conditions and could relate to the need for excitement seeking.

In group 5, two variables correlated positively (0.229 and 0.254) and two variables correlated negatively (−0.229 and −0.240).

Overall, the patterns for extraversion may be that extraverts seek excitement in general and they may see phenomena like strange lights, movement, and threats as exciting. High extraversion scorers might move around the game to a large extent in order to try to locate exciting phenomena. Our game may be seen as having low levels of excitement in general because we did not choose to include the overt violence often found in video games.

Agreeableness

In group 5, eight variables correlated positively (between 0.222 and 0.284) and five variables correlated negatively (between −0.234 and −0.304). The variables that agreeableness correlates to are those related to warning the villagers about the poisoned water in the lake. The remaining variables are common courtesy variables such as saying “thank you” and greeting others in a friendly manner. The negative correlations show that high agreeableness scorers tend to avoid unfriendly or mean remarks.

Overall, agreeableness shows the highest correlations after openness. Agreeableness also shows a large number of correlations at the $p < 0.01$ level. These highly significant relationships are conversation variables related to friendly interaction. It seems that high agreeableness scorers tend to act kindly towards others to some extent.

Neuroticism

In group 3 we see that one variable correlated negatively with neuroticism (−0.229). The high neuroticism scores tend to have more conversations in the village area.

In group 5 we see that seven variables correlated positively (between 0.226 and 0.265) and three variables correlated negatively (between −0.233 and −0.310).

Overall, we see that high neuroticism scorers (1) tend to take more time to finish the game than low scorers. We have no explanation for the correlations related to high neuroticism scores.

5.5.2 The effects of openness

When examining the results of Section 5.4 we may conclude that openness is the most influential variable for the overall game behaviour effects. High scorers tend to finish the game more often and show less conversation and movement-triggers. The values of movement trigger variables are increased by moving around the world. Moving around could be related to explorative behaviour of the game world or to goal-directed behaviour. Openness is often linked to curiosity and the willingness to try new things. Surprisingly, the results show that openness has negative correlations to many of both the individual variables as well as the pooled variables. At first glance, this effect seems counter-intuitive. We would expect individuals with high openness to be interested in the various aspects of the game that can be explored, but when we inspect the data we see the opposite effect. We present two possible explanations for this effect: (1) high scorers are mostly interested in novel situations and thus hurry through the game looking for variation while experiencing the various types of situations briefly and (2) the module we built resembles commercial games so closely that it may present a familiar experience to the player. Over-familiarity could cause high-openness individuals to seek their novelty elsewhere.

5.5.3 Interpreting the correlations

In this experiment, we examine the correlations between an *original choice* and its alternatives. Positive correlations in conversations can have two explanations: (1) players opt for an *original choice* because it is more attractive than the other available choices or (2) the other choices were in some way less attractive compared to the *original choice*. With both positive and negative correlation, this means that (a) the *original choice* was less desirable than the other choices, or (b) some aspect of the alternative choices made the *original choice* more important. These distinctions are relevant in both cases because explanation (a) implies that the reason for actively choosing or avoiding a choice is inherent to the correlation variable, while explanation (b) implies that the cause of the variation lies in other variables but is expressed in the correlation.

For example, if a player has two conversation options, X and Y, and he chooses option X this could mean at least one of five things, (1) option X is experienced as positive and option Y is experienced as positive but slightly less so than option X, (2) option X is experienced as positive and option Y as negative, (3) option X is experienced as positive and option Y as neutral, (4) option X is experienced as neutral and option Y as negative, and (5) option X is experienced as negative but option Y is more negative. Interpretation of the results should be considered very carefully in order to avoid reaching an incorrect conclusion.

We note that in interpreting the correlations, the risks described in Subsection 4.5.2 should be kept in mind.

5.5.4 The modelling of extraversion

The linear regression analysis shows that we are capable of fitting a linear model to our game data in order to predict four of the five OCEAN variables accurately. The variable with the least variance explained was extraversion. This result was quite unexpected since early research into factor models of personality revealed extraversion as one of the first traits to be discovered (Digman, 1990). According to Digman, extraversion was found to be prevalent in many facets of life. We expected that extraversion would be present in game behaviour based on the results from Chapter 4, but we did not know if the other traits could be modelled successfully. Two possible explanations for the lower predictability of extraversion in the behaviour that we have measured might be: (1) the behaviour we have measured was unsuitable for extraversion measurement, or (2) extraversion may be less prevalent in game behaviour.

5.5.5 Interpreting regression results

At low numbers of participants in comparison to the number of predictors, regression runs the risk of overfitting. If a regression model overfits, it overestimates the amount of variance it explaines. Field (2009) suggests the following rule. Per estimator variable in the model there should be at least 10 participants in the experiment.

(Field, 2009) proposes three solutions for the overestimation problem in regression: (1) Limit the number of predictors used in the model to a maximum of 1 per 10 participants, (2) increase the number of participants to 10 per predictor that is required in the model, (3) reduce the number of predictors by summing the z-scores of predictors that measure the same trait (an investigation using Cronbach's Alpha can be used for this).

5.5.6 Effect sizes and significance

Our correlations do not often have large effect sizes (correlation of 0.5 or higher) but they are significant. One reason which we suspect is causing the lack of large correlation effect sizes is the fact that we did not fine-tune our game as is done with a personality test. There may be dozens of factors influencing behaviour in a natural setting. Because of this it is reasonable to expect low correlations when looking at any single factor. As stated before in Section 4.4, correlations between 0.1 and 0.5 are, when significant, reasonable results. We wished to see whether personality effects can be found in games similar to those normally played and we have succeeded. The research in Chapter 4 indicated that trait prediction can be improved by putting situations in the game that are optimally suited to each trait's expression (cf. van Lankveld et al., 2011a).

5.5.7 Similarities between conversations and multiple-choice questions

When considering the resemblance between questionnaires and the conversations in our game, the argument could be made that our conversation items are multiple choice questions and that they are therefore the same as the multiple choice items in personality questionnaires. Some readers may make the argument that the conversations might as well be done via an offline questionnaire.

Conversation choices in the game are multiple choice but they are not the same as the multiple choice items in a personality questionnaire. See the conversation example in Subsection 5.3.2 for a clarification of a multiple choice item in our game. Items in a questionnaire are descriptive statements on which the participant has to rate himself. In contrast, conversation choices in the game are responses to NPC statements. These statements are not descriptive statements about the player. The player expresses his personality through his behaviour in the game rather than describing himself.

5.6 Chapter conclusions

In our research we tested 80 participants on 260 game behaviour variables and on the NEO-PI-R personality questionnaire.

From our results we may conclude that personality effects on game behaviour exist for all five traits of the FFM. We found these effects when we performed linear regression analysis and when we performed detailed correlation analysis of the game behaviour variables to scores on a personality questionnaire. Therefore we may conclude that we are able to produce accurate estimates of a participant's personality based on the game variables.

We investigated RQ3: *To what extent can we use games in order to create a full personality profile automatically?* In our experiments we found that all the elements are present for the creation of a personality profile.

This investigation examined the effects of personality in only one game of the role-playing game type. In the next chapter we investigate how a personality assessment interview compares to the NEO-PI-R questionnaire. This investigation should provide clarity regarding the position that observational studies provide more insight into personality than questionnaires do. We also investigate whether a theory-driven model can provide more insight into personality in games than the data-driven model used in this chapter.

Chapter 6

Theory-driven personality in games

In this chapter we investigate RQ4: *To what extent does a theory-driven model explain personality in games?* We can use existing theories in personality literature to create a model of personality in games (i.e., a theory-driven model). We examine to what extent a theory-driven model which is supplied with gameplay data correlates to the 'Big-Five' personality traits.

In Section 6.1 we provide an introduction to the research in this chapter. In Section 6.2 we explain our theory-driven model in detail. In Section 6.3 discuss the methods used in conducting our investigation. In Section 6.4 we give the results of our investigation. In Section 6.5 we discuss our results. In Section 6.6 we give conclusions to our research.

6.1 Alternatives to data-driven models

The results from the previous chapter show that improvements could be made to the explanatory value of data-driven model results. Less than 20% of the data-driven game variables correlated to the 'Big-Five' personality factors and there is a possible overfitting effect in the linear regression model results. The model variables that do correlate are hard to interpret because a clear pattern of effects that can be described as personality traits is lacking. The variables that are incorporated in the linear regression model for each specific personality trait are not clearly related and they do not coincide with the variables found in the correlation analysis. In summary, (1) improvements are necessary before the model can be used in novel situations and (2) a clear conclusion based on the correlation analysis is lacking.

One of the problems that clearly arises in Chapter 5 is that the ratio of variables to the number of test participants is unfavourable. In parallel, significance thresholds that are more strict are preferable when dealing with a large number of correlation analyses. A power analysis shows that for 260 variabeles a minimum of 1600 participants is necesary in order to improve the chance of finding statistical effects up to 80%. The 80 participants used in the previous chapter leave too much chance to miss effects and to accept results that are merely statistical noise.

We propose to reduce of the number of variables in the model as a solution to the problem described above. By using a theory-driven basis we may reduce the 260 data-driven variables to a maximum of three variables per personality trait. Subsequent analysis will demonstrate whether such a theory-driven model is effective in handling the issue while maintaining predictive validity.

6.2 A theory-driven model

In order to create a theory-driven model we re-examined the literature of Chapter 5. Based on the litature, we attempted to find personality-related behaviours that can be found in games. For each of the five personality traits, the available literature provides some behavioural equivalent which can be examined by looking at game behaviour. Below, we provide a list of the behaviours that we expect to see in video games based on the behaviour descriptions that personality theory provides for each trait.

- *Openness*: openness variables revolve around the expression of curiosity and the wish for new information and stimuli. Individuals with a high openness score are expected to look for variation in life, they tend to break set habits. In other words, they look for novelty and are quick to trade familiar situations in life in favour of original substance. In games, we expect high openness scorers yearn to experience many things at least once and not stick to habits for too long.
  We expect high scorers (1) to explore many conversation options but not to return to them often, and (2) to move to many different locations but not to return to them often.

- *Conscientiousness*: conscientiousness coincides with planning, keeping to social protocol, being reliable and fully completing tasks. In games, we expect high conscientiousness scorers to be more likely to complete our game scenario in all details. We also expect them to honour agreements and promises to NPCs more often than low conscientiousness scorers.
  We expect high scorers (1) to be more likely to finish all conversations in the main story, (2) to be more likely to complete movement to all locations in the main story, and (3) to be likely to honour all agreements they make in the game.

- *Extraversion*: extraverts look for challenge and excitement. In games, we expect high extraversion scorers to display much proactive search behaviour. We also expect high scorers to converse and move around more than low scorers.
  We expect high scorers (1) to have a higher number of conversation choices overall and (2) to move around more in general.

- *Agreeableness*: individuals with high agreeableness scores are expected to be friendly and helpful. In games, we expect high agreeableness scores to choose more friendly and more considerate conversation options as well as avoid harmful, brash or unfriendly conversation options.

We expect high scorers to choose a higher number of positive and friendly conversation options while avoiding unfriendly or hurtful conversation options.

- *Neuroticism*: people that score high on neuroticism are expected to be insecure and feel negative emotions more often than low scorers. In games, we expect high scorers to take more time to complete tasks due to their uncertainty and therefore we expect them to take longer in completing the game. We expect high scorers to ask for more information in order to improve the knowledge of the situation and to gain more certainty.
  We expect high scorers (1) to proceed slower (more cautiously) through the game, (2) to choose more optional conversation choices, and (3) to move to a higher number of optional locations.

We have considered the raw data that we have collected in Chapter 5 in order to find representations in game behaviour for the behaviours described above. We have formulated a list of behavioural criteria that allow for the measurement of the described behaviours. The creation of behavioural criteria variables requires expert knowledge and is therefore more labour intensive than the creation of the data-driven model variables described in the previous chapter.

Definition 6.1 (Behavioural criteria) Behavioural criteria are defined as measureable variables that result from the operationalisation of a psychological theory.

Definition 6.2 (Operationalisation) Operationalisation is defined as the process of creating measureable variables from a theory for the purpose of a precise measurement of the theory's concepts.

Together, the behavioural criteria variables form the theory-driven model of game behaviours. Our theory-driven model consists of eleven variables, two for openness, three for conscientiousness, two for extraversion, one for agreeableness, and three for neuroticism. The variables in our model are all made up of raw gameplay data. We named these new variables "new criteria" (NC for short). The new critetia for openness are referred to as NCO, for conscientiousness we refer to them as NCC, et cetera for the other traits. The list of constructed variables is as follows.

For the openness variables NCO1 and NCO2 we chose to investigate game variables activated exactly once because individuals with high openness crave variation and are therefore expected to investigate new content often. They also grow tired of known content quickly and are thus likely to spend only little time with the content. In Chapter 5 we found that high openness scorers proceeded more rapidly through the game.

For the conscientiousness variables NCC1 and NCC2 we expect that high scoring individuals are more likely to finish the tasks they are currently working on. The game is implicitly a task that the players are working on, so we expect high scorers to be more likely to finish all the points of the main story. For NCC3 we expect high conscientiousness scorers to be more reliable and therefore to keep more agreements or promises made in the game. Contrary to our expectations in this chapter, in Chapter 5 we found that high conscientiousness scorers invest more time in exploring sidequests.

Table 6.1: Theory-driven model variables.

Name	Implementation
	Openness
NCO1	+1 per conversation chosen exactly once
NCO2	+1 per location moved to exactly once
	Conscientiousness
NCC1	+1 per conversation in the main story
NCC2	+1 per movement in a main story location
NCC3	+1 per honoured agreement, -1 per agreement not honoured
	Extraversion
NCE1	+1 per conversation
NCE2	+1 per moved location
	Agreeableness
NCA1	+1 friendly conversation choice, -1 per unfriendly conversation choice
	Neuroticism
NCN1	total time needed to finish the game
NCN2	+1 per optional conversation
NCN3	+1 per option moved location

For the extraversion variables NCE1 and NCE2 we expect high scoring individuals to be explorative and seek excitement and social situations. In Chapter 5 high extraversion scorers tended to move around more in specific areas in the game such as the dream area.

For the agreeableness variable NCA1 we expect high scorers to be more friendly and to avoid unfriendly, unreliable, or untruthful conversation choices. In Chapter 5 high agreeableness scorers tended to choose the friendly conversation options to warn NPC characters. They also showed other friendly conversation preferences. In the game, we find fourteen possible friendly conversation options and nine possible unfriendly conversation options that are available for the player's choice.

For neuroticism variable NCN1 we expect higher scorers to be more insecure and therefore require a greater amount of time in order to check things resulting in a higher total time needed to finish the game. For NCN2 and NCN3 we expect high scorers to move around more and talk more in optional parts of the game to explore and satisfy their need for security. In Chapter 5 we found that high scorers tended to converse more in the village area and that they take longer to finish the game. Each of the theory-driven model variables was calculated using the data collected in the research of Chapter 5.

6.3 Experimental setup

As explained above, the theory-driven model was constructed using the collected raw data from Chapter 5. We have used SPSS for the creation of the eleven behavioural-criteria variables. For the creation of NCO1 and NCO2 we needed to extend the original data collected in Chapter 5 with variables measuring the conversations and movement locations with a value of exactly one. The new theory-driven dataset was calculated using the 80 participants of Chapter 5. The new dataset was built up from the 217 raw variables which contained movement, conversation, and general data.

The dataset contained the behavioural data from the game NWN in the scenario 'the poisoned lake'. The data collected in the original research can be found in Chapter 5 and will not be repeated here. The NEO-PI-R personality data that we use in this chapter are also the same data collected in Chapter 5 and will also not be repeated here.

6.4 Results

The descriptive statistics of the personality traits are identical to those found in Chapter 5 and can be found in Table 5.2. The descriptive statistics for the new theory-driven model variables are given in Table 6.2.

Table 6.2: Descriptive statistics.

Variable	N	Minimum	Maximum	M	SD
NCO1	80	19	46	34.24	6.311
NCO2	80	9	28	17.55	3.482
NCC1	80	6	19	13.29	2.571
NCC2	80	7	14	9.96	1.563
NCC3	80	-1	2	.48	.779
NCE1	80	19	69	42.89	10.920
NCE2	80	11	19	15.39	1.717
NCA1	80	0	11	4.64	2.616
NCN1	80	737	4328	1846.89	860.448
NCN2	80	38	92	67.15	11.316
NCN3	80	21	47	37.08	5.891

After both a visual check of the distribution and a skewness analysis we see that all the variables in Table 6.2 have a skewness much smaller than the 1.96 threshold set in Field (2009) and are considered to follow a normal distribution.

In order to examine the relationships between the variables of the theory-driven model and personality we have performed a correlation analysis for each of the five personality traits with all of the theory-driven variables. The results of the correlation analysis is given in Table 6.3. We see the five personality traits in the top row and we see the eleven theory-driven variables in the leftmost column. The values displayed are Pearson's correlations. We remark that only significant correlations are displayed. Correlaties with a * are significant at the $p < 0.05$ level and correlations with a ** are significant at the $p < 0.01$ level.

In Table 6.3 we see one ** correlation for agreeableness, three * correlations for neuroticism, and no other significant correlations. The values of the significant correlations fall between $r = .248$ and $r = .304$.

Table 6.3: Correlations.

	Openness	Conscientiousness	Extraversion	Agreeableness	Neuroticism
NCO1					
NCO2					
NCC1					
NCC2					-.241*
NCC3					
NCE1					
NCE2					
NCA1				.304**	
NCN1					.248*
NCN2					
NCN3					-.269*

6.5 Discussion

The significance of the correlation for the NCA1 variable and agreeableness is the highest of the correlations found, a chance of one in 100 that we have found this result by coincidence. Interpretation of this result is straightforward, participants that score higher on agreeableness are more likely to choose friendly conversation options (+1 to NCA1) and more likely to avoid unfriendly conversation options (-1 to NCA1). There are fourteen possible friendly conversation options and nine possible unfriendly conversation options. However, since each option can be chosen multiple times, the possible range of the scores is unlimited. The actual range of the collected data was between zero and eleven. If a random distribution of friendly and unfriendly conversation choices would occur we would expect a mean of 1, the actual mean was 4.64.

Neuroticism correlates positively to NCN1 and negatively to NCN3 and NCC2. A priori, we expected only correlations between neuroticism and NCN variables. The correlation between neuroticism and NCC2 was unexpected. NCC2 represents the total number of locations that need to be visited in the process of completing the main storyline of the game.

Upon closer inspection, the total time needed to finish the game could have a relationship to the total number of visited locations. If the player finishes the game, he has visited all the locations of the main storyline. It seems logical that, if a player is slower (i.e., has a higher total time) the chance of the participant not visiting all the locations of the main story increases. An extra correlation analysis between NCC2 and NCN1 was done. This analysis shows that there is no significant relationship between both variables (r = .135, p = .233) so the logical reasoning described above does not apply. Additionally, the data shows that even for the slower players there is sufficient time to visit all main storyline locations.

As stated earlier, neuroticism correlates positively with the total amount of play time. This means that a higher neuroticism score has an increased probability of coinciding with a higher score for total time. We cannot pinpoint an exact cause for this effect in the current investigation although the observations do fit with the theoretical explanations described in

Section 6.2. It is also notable that higher neuroticism coincides with a higher NCN3 score, the total amount of visited optional locations.

6.6 Chapter conclusions

In this chapter we investigate RQ4: *To what extent does a theory-driven model explain personality in games?* We have conceptualised and operationalised a theory-driven model for personality in games. We have also tested the theory-driven model. The model consists of eleven variables in contrast to the 260 variables of the data-driven model described in Chapter 5. The tests of the theory-driven model reveal correlations between model variables and personality traits for four of the eleven model variables. The correlations found numbered three for the neuroticism trait and one for the agreeableness trait. So, on the basis of these findings, we are inclined to conclude provisionally and tentatively that (1) statistical relationships exist between our theory-driven variables and the five personality traits and that (2) the percentage of correlations for the theory-driven model surpasses the percentage of correlations found for the data-driven model.

The words 'provisionally' and 'tentatively' ask for some foreseeable critique. We are open to that and will start with an important question that holds for both the theory-driven as well as for the data-driven model. The question is whether the results are maintained across different situations and across different experiments. Personality is considered to be stable across situations. Therefore, to assess the generalisability of both models we will attempt to validate both of them on a new game in the next chapter. We also attempt to validate the concept of personality in games by comparing the predictive validity of game behaviour and personality questionnaire scores to real world responses in an interview. The results of this interview study can be found in Appendix N.

Chapter 7

Validating personality in games

In this chapter we investigate RQ5: *To what extent can our models of personality in games be validated in different games?* To answer this question, we investigate the generalisability of our proposed model of player profiling. If it turns out to be generalisable, we then test our model on a commercially available and critically acclaimed game: FALLOUT 3[1].

In Section 7.1 we provide an introduction to the generalisability of our model and to personality profiling in FALLOUT 3. In Section 7.2 we present the precise characteristics of the game used in this experiment. In Section 7.3 we provide the experimental setup. In Section 7.4 we present the results, which are discussed in Section 7.5. In Section 7.6 we give our conclusions.

7.1 Profiling by using a commercial game

We attempt to apply automated personality profiling techniques to a commercially available video game in order to (1) assess the generalisability of our model, (2) provide further validation for our profiling technique, (3) obtain a more extensive understanding of the available potential for the expression of personality in a commercial video game. In Chapter 5 our setting was tuned to provide a broad range of responses. Now, we aim at obtaining a profile without affecting the player's experience of the commercial game. Since we created the scenario ourselves we ran the risk of introducing our biased view into the scenario. Such a bias may reduce our game's similarity to a commercial game experience. In summary, the potential bias then affects the ecological validity of our approach. Ecological validity relates to the experimental setting being equal to a "real-life" setting. As a consequence, the result is that the behaviour in the experimental setting is the same as the behaviour in the natural setting. The commercially available game application is meant to reduce the risk of bias and to increase ecological validity.

Definition 7.1 (Ecological validity) Ecological validity measures the similarity of an individual's behaviour in an experimental environment compared to his behaviour in real life.

[1]Published by *Bethesda Softworks* (2008)

If the ecological validity is high, behaviour in the experiment is similar to behaviour in real life.

7.2 The game: vault 101

The purpose of current research is to investigate the possibilities of a commercial computer game to create personality profiles in the same manner as we did with a Neverwinter Nights module in Chapter 5. In the current research, the commercial game Fallout 3 is used. This section explains the choice for the scenario "vault 101" in the game Fallout 3 and gives information about the story, the setting, and the controls of the game.

7.2.1 General game information

Fallout 3 is an action role-playing/adventure game; it is the fourth game in the Fallout series, which can be played in both first- and third-person perspective. Fallout 3 is chosen because it is a modern game which contains both movement and conversation behaviour. It has an introduction which does not contain deadly violence, and which presents a variety of situations to the player. Here we remark that only the introduction of the game is played by our participants.

7.2.2 Story

The scenario of Fallout 3 contains an adventure type of story, which is common for role-playing games. The story consists of an introduction and a central game. Fallout 3 takes place in the year 2277, 200 years after a nuclear holocaust, which devastated the world. In the introduction the player character grows up in a nuclear shelter (named vault 101).

The introduction, which is the only part that we used for our experiment, has four phases.

1. birth: During this phase the player of the game can decide which gender he wants to be and what he wants to look like.
2. 1 year old: This is also a training phase in which the player can practice the controls by walking around in a room.
3. 10th birthday party: At the party the player can talk to 12 people and he receives presents.
4. 16 years old: During this phase the character has to do a career test named G.O.A.T., which stands for Generalized Occupational Aptitude Test.

After the first four phases of the introduction there is a fifth phase which we do not use because of time constraints and because the game becomes much more violent in the fifth phase. In the fifth phase, the player character is nineteen years old. He wakes up in the vault and has to escape to the world outside the shelter where the central game starts. No further information is given about this last part of the story, because the experiment ends after the fourth phase. An overview of the characters encountered in the game is given in Table 7.1.

Table 7.1: Names and descriptions of the NPCs.

Name	Description
Amata	A friend of the player character.
Beatrice	An acquaintance of the player character.
Butch	The leader of a gang called "Tunnel Snakes".
Dad (James)	The father of the player character.
Jonas	A scientist and friend of dad.
Mr. Brotch	The teacher who is responsible for the G.O.A.T.
Officer Gomez	A friendly security guard.
Old Lady Palmer	A friendly old lady.
Overseer	The boss of the vault and the father of Amata.
Stanley	A friendly engineer.
Paul Hannon	A boy, starts friendly but is a member of Butch's gang later on.
Wally Mack	A friend of Butch, who does not like the player character.

7.2.3 Setting

Fallout 3 takes place in a nuclear shelter, vault 101. The shelter looks cold and drab because the walls consist of metal. There are tunnels and doors, some of which can be opened. When opening doors the player can move to other rooms and tunnels. The player has to perform specific actions to be able to proceed through the shelter which results in a linear gameplay experience (i.e., the player must follow the game's plot points one-by-one).

7.2.4 Controls

The interaction between the player and Fallout 3 happens by mouse and keyboard. Only the most important controls are explained here. The keys 'W','S', 'A', and 'D' are used as arrows to move forward, backward, left, and right. This is the standard control scheme for most contemporary 3D games. The key 'E' is used to start conversations with other characters or for actions such as grabbing things or opening doors to move to another area/room. Whenever 'E' can be used, an indication is shown on the screen. The 'space' bar can be used to jump. The mouse is used to control the camera view, so that the player can look everywhere around him. When the player starts a conversation with another character, the conversation happens via a menu with several possible responses. To choose a response, the player clicks with the left mouse button on one of the responses. The number of possible responses varies per conversation.

7.3 Experimental setup

In the experiment, participants play the game Fallout 3. While playing, the in-game behaviour is logged in a file. After playing the game for a maximum of 45 minutes, the participants have to complete two questionnaires; one concerns control factors, and the other one is a personality test. The personality test used in this experiment is the NEO-FFI. We use the NEO-FFI because it requires only 15 minutes to complete (versus 45 minutes for the NEO-PI-R) (see Hoekstra et al., 2007). We do not perform a test on order effects because the research in Chapter 4 showed that no order effects were to be expected.

In Subsection 7.3.1 the participants and the experimental procedure are discussed. In Subsection 7.3.2 we explain which data-driven variables are used, and in Subsection 7.3.3 where theory-driven variables are used. In Subsection 7.3.4 we describe the questionnaires. In Subsection 7.3.5 we present our analysis.

7.3.1 Participants and procedure

We recruited 36 participants via a student pool. They all received research credits for participating in the experiments. The age of the participants ranged from 18 to 27. The mean age of the participants was 21.5. 18 participants were male and 17 were female (for one person the gender is unknown). Only students who had never played the game FALLOUT 3 before were allowed to participate.

The experiments were always executed in the same room in one of the buildings of Tilburg University, which was made available for these experiments. The room was an office room, but tidied up so that there were no distractions. The participants were seated at a desk with a computer on it. Effort has been put in making the conditions equal for all participants, so that the environment could not influence the results of the experiments. When the participants entered the experiment room, they received a sheet with instructions for the game. The instructions were written in Dutch, because all the participants had the Dutch language at their command (it was either their native language or their second language). The instructions are given in Appendix I. The instructions explained the most important controls and lead the participants through the first two phases. After reading the instructions the participants played the game. They were able to consult the instructions while playing the game. During the first two phases the participants had the opportunity to ask questions. During the third and fourth phase the participants were not allowed to ask any questions, even though the experimenter stayed in the room in case of emergencies. In that time, the experimenter did not observe the participants actively. The experimenter sat at a distance of approximately 3 meters, and was not able to look at the computer screen from that position.

The participants had about 45 minutes to complete all phases. After the fourth phase a pop-up appeared in which the participants were asked to stop playing and to call the experimenter. If the participants did not manage to complete all phases in 45 minutes the experimenter asked them to stop playing. After playing the game, the participants had to fill in two questionnaires. The results of the questionnaires were anonymous, because the participants did not have to give their names.

Four of the 36 participants could not complete the game in 45 minutes. Therefore, their experiments had to be interrupted. The participants had reached the final phase of the experiment but they had not started with the G.O.A.T., being the last part of the game. The activation of movement and conversation variables prior to the G.O.A.T. are independent of the scores on the G.O.A.T. and are therefore included in the theory-driven analysis. Here, we note that (1) the G.O.A.T. variables are slightly less reliable, and (2) not reaching the G.O.A.T. has no influence on the data-driven variables of Subsection 7.3.2.

7.3.2 Data-driven variables

The Garden of Eden Creation Kit (in brief: GECK) is used to edit and create data for using FALLOUT 3 for the current research. The created data is logged in a log file, and used to analyse the behaviour. In total, 165 variables were created, 107 variables recorded the participants conversation behaviour, 20 variables recorded movement behaviour, 37 variables recorded the choices within the G.O.A.T., and one variable recorded whether the participant became involved into a physical fight with NPCs. Only the behaviour in the third and fourth phase of the game is recorded. There are two types of variables (1) pooled variables and (2) unpooled variables. All variables can be found in Appendix L. The variables are divided into five groups.

- Group 1 contains two pooled variables; a total conversation variable and a total movement variable. In the total conversation variable, all conversation variables are added up. Similarly, all movement variables are added up.
- Group 2 contains seven pooled movement variables. There are pooled movement variables for five sub-areas in the game. The two phases themselves had a pooled movement variable as well. In this way the two phases can be compared, and it can be checked which areas are responsible for certain results within a phase.
- Group 3 contains two pooled conversation variables. More specifically, there is one variable for all conversations which might take place during the phase in which the player character is ten years old and one variable for the phase in which the player character is sixteen years old.
- Group 4 contains a pooled conversation variable per NPC; sixteen variables in total. The NPCs which had a role in both the ten year old as well as the sixteen year old phase, are taken separately.
- Group 5 contains all 165 unpooled variables.

After the composition of these five groups, there are 192 variables in total. This total number consists of 165 unpooled variables and 27 pooled variables. For the sake of clarity we have included a complete list of the conversation choices in group 5 and their variable names in Appendix K.

Pooled variables

Pooled variables are calculated in the same way as in Chapter 5. By summing the values of several unpooled variables their higher order effects become clear. The assumption is made that there might be average effects of conversation or movement that only appears across the entire areas of game behaviour. The pooled variables collect the values of all unique variables per phase, per NPC, and for all phases together (i.e., the two phases which the participant can play). In our case the following happened: all values of the variables are summed to form the pooled variables.

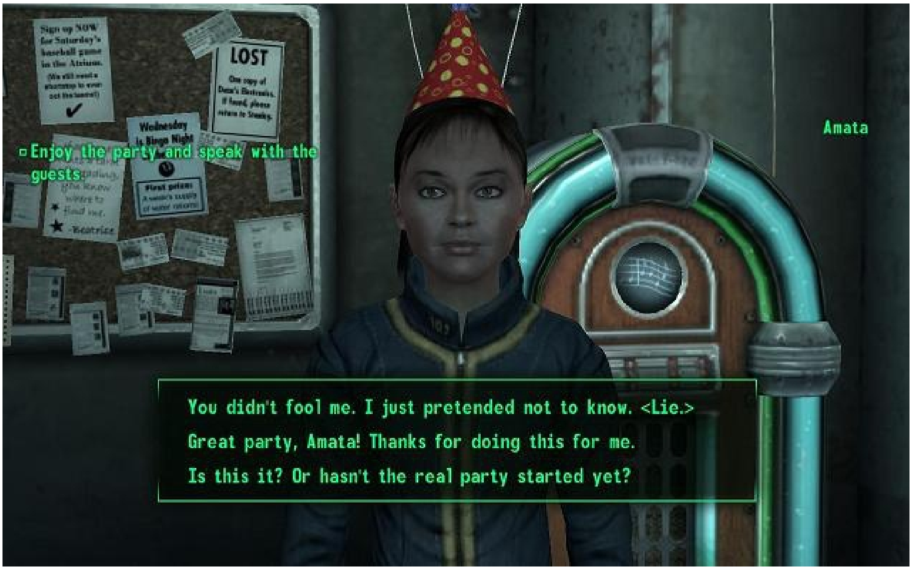

Figure 7.1: A screenshot of a conversation in FALLOUT 3.

Unpooled variables

There were four types of unpooled variables, (1) conversation, (2) movement, (3) answers on a test in the game called G.O.A.T, and (4) whether the player became involved in a fight with a gang (called "Tunnel Snakes"). Each variable recorded how many times the corresponding behaviour was performed. For conversation and movement it is possible to perform the same behaviour more than once. The choices of response within conversations were recorded by the conversation variables. Figure 7.1 shows an example of three conversation options. Each movement variable recorded the total number of movement behaviours. Every time the participants (1) chose a certain option within a conversation, (2) entered an area of the game monitored by a movement trigger, (3) chose an answer for the G.O.A.T., or (4) became involved in a fight with the gang, the corresponding variable increased by one point. The movement variables were placed in front of all inaccessible doors and at crossings of corridors. All variables in the game were natural and unlimited (In practice the largest range is 28).

7.3.3 Theory-driven variables

The theory-driven variables were created based on our theoretical personality behaviours composed in Chapter 6. In contrast to the research of Chapter 6, in Chapter 7 a total of eight variables was composed. The values of these variables were computed after the dataset for the data-driven variables was collected and their values were known. The variables that were created are given in Table 7.2.

Table 7.2: Theory-driven model variables.

Name	Implementation
	Openness
NCO2	+1 per location moved to exactly once
	Conscientiousness
NCC1	+1 per conversation in the main story
NCC2	+1 per movement in a main story location
	Extraversion
NCE1	+1 per conversation
NCE2	+1 per moved location
	Agreeableness
NCA1	+1 friendly conversation choice, -1 per unfriendly conversation choice
	Neuroticism
NCN2	+1 per optional conversation
NCN3	+1 per option moved location

The variables were created in a similar fashion to the variables in Chapter 6. There are three variables not included in this list: NCO1, NCC3, and NCN1. NCO1 was not included because in the game of FALLOUT 3 conversation options could be chosen only once. Therefore, a variable measuring the conversation options that were chosen exactly once is redundant. The variable NCC3 was not included because there were no opportunities to make promises or agreements in the story of FALLOUT 3. Finally, the variable NCN1 was omitted incidentally. It was not included in the raw variables at the time of the experiment.

Moreover, the precise implementation of the variables in this chapter differs from the implementation of the variables in Chapter 6. As stated above, the game was originally set up so that conversation options could not be revisited. Furthermore, since the game was not in the outdoors there was only one main route that players could follow. Therefore, the creation of main-storyline-movement variables was deemed as “mostly unnecessary”. The exceptions were two movement variables, trigger04 and trigger15. These movement variables were included to see if the player would backtrack to visit earlier locations in the “party” and “shooting” parts of the 10 year old phase.

7.3.4 Questionnaires

For the current research two questionnaires were used. One questionnaire is based on a questionnaire used by Schreurs (2009). This questionnaire is given in Appendix J. The second questionnaire was the NEO-FFI, which is an abridged version of the NEO-PI-R. The NEO-PI-R is described in more detail in Subsection 2.5.3. The NEO-FFI is based on the same five traits as the NEO-PI-R. The NEO-FFI has fewer questions, namely twelve per trait. Therefore it can be processed in approximately 15 minutes. Because the study by Schreurs (2009) showed that there were no order effects concerning whether the participants started with the questionnaires or the game, in the current research we did not take the occurrance of such effects into account.

7.3.5 Analysis

The outcomes of the questions per trait were totalled, and then normalized. The outcomes were normalized into stanine scores (scores from 1 to 9) conforming to the norms table in Hoekstra et al. (2007) used for participants in scientific research.

We analysed the bivariate correlations. A bivariate correlation measures the relationship between variables and the strength of this relationship. In the current research, the relationship between the game items and the traits were measured. This was also done for the control factors of the questionnaire about demographics and experiences. So, our analysis could discover whether some outcomes were due to the control factors, the traits, or both. The correlations resulted in several positive and negative significant outcomes. Correlations were considered to be significant when $p < .05$. For an interpretation of the effect sizes, we refer to Section 4.4. In addition to the correlation analysis, we performed a linear regression analysis in order to see whether the OCEAN traits could be predicted using the pooled and unpooled game variables.

The significant outcomes of the bivariate correlations are used for linear regression. The traits are used as independent variables in every analysis. The control factors are only added to the independent variables when a significant correlation is found. In the dependent variable the game items (pooled and unpooled) are placed one at a time. This is done using the stepwise method, where attention is paid to the R^2. The R^2 is used to measure the effect size (see also Chapter 5).

For the theory-driven model variables we analysed their descriptives followed by a Pearson's correlation analysis between each individual variable and each individual trait of the OCEAN personality model.

7.4 Results

The results are partitioned into two subsections. In Subsection 7.4.1 we analyse the results of the data-driven model validation. In Subsection 7.4.2 we analyse the results of the theory-driven model validation.

7.4.1 Data-driven results

The normalised outcomes of the NEO-FFI range from 1 to 9 (stanine scores) and are given in Table 7.3. Our measured traits show slight deviations from the expected stanine values in the population. The ranges of openness and extraversion have higher minima than expected while conscientiousness, agreeableness, and neuroticism do not reach the highest possible scores. However, the means and the standard errors found in the current research are similar to the scores as found in the standardized NEO-FFI.

Below, we compare the results of this investigation to the results found in Chapter 5. Table 5.3 shows predictions of all personality traits with a certainty above 55% with the exception of extraversion (35.1% of the variance is explained). The results of the current investigation are shown in Table 7.4. The details on the interpretation of the linear regresion results are given in Section 5.4. We note that in interpreting the current regression results, the risks described in Subsection 5.5.5 should be kept in mind. We see that for the current

Table 7.3: Descriptive statistics of NEO-FFI scores.

	Minimum	Maximum	M	SD
Openness	3	9	5.81	1.60
Conscientiousness	1	7	4.03	1.54
Extraversion	2	9	5.86	1.87
Agreeableness	1	8	4.17	1.99
Neuroticism	1	7	4.53	1.63

Table 7.4: Stepwise linear regression analysis of traits and game variables.

Trait	R^2	Variables in model
Openness	.982	15
Conscientiousness	.486	4
Extraversion	.837	8
Agreeableness	.752	6
Neuroticism	1.000	23

investigations, we are able to predict openness and neuroticism with certainty above 95%, extraversion and agreeableness with certainty between 75% and 85%, and conscientiousness with certainty below 50%.

In order to examine the results of our new investigation more closely, the one-tailed correlations corresponding to the linear regression analysis are examined in detail. Table 7.5 contains the total number of positive and negative significant correlations per group of variables. We include only significant correlations in the table; an empty cell means there is no significant correlation. As can be seen in Table 7.5 significant correlations are found for every group of game variables.

Each of the five personality traits investigated in this chapter has at least one correlating variable. This means that significant effects are present both in conversation and in movement behaviour in the game. The largest number of correlations are found for the traits extraversion and neuroticism. Agreeableness has significant correlations in four out of five groups. It can be observed that there are a total of 101 correlations with a total of 40 positive correlations and 61 negative correlations.

Furthermore, a total of 87 of the 192 game variables show significant correlations with one or more personality traits. This number differs from the 101 variables included in Table 7.5 because some variables correlate to more than one trait and are thus overlapping. It means that the behaviour measured by the variable is significantly influenced by multiple traits. The current number (87) is 45% percent of all the game variables. The results are based on the complete list of significant correlations, which can be found in Appendix M. Table 7.6 shows the minimum and maximum correlations found for each group. We note that in interpreting the correlations, the risks described in Subsection 4.5.2 should be kept in mind.

Table 7.5: Positive and negative correlations per variable group.

	Group 1		Group 2		Group 3		Group 4		Group 5		Total	
	Pos	Neg	Pos	Neg	Pos	Neg	Pos	Neg	Pos	Neg	Pos	Neg
O							1		12	7	13	7
C								1	3	9	4	10
E							1		16	6	17	7
A		1		3			1		4	11	5	15
N		2				1		3	1	16	1	22
Total		3		3		1	3	4	36	49	40	61

The letters OCEAN (left column) stand for their respective traits.

7.4.2 Theory-driven results

Table 7.7 shows the descriptives of the theory-driven variables in this chapter. Please note that smaller distances from zero are referred to as lowest while larger distances from zero (both positive and negative) are referred to as highest.

We present the results of the correlation analysis for the theory-driven variables and the OCEAN personality traits in Table 7.8. The table shows two significant correlations at the $p < 0.01$ level: agreeableness - NCE2 ($r = -.444$) and agreeableness - NCN3 ($r = -.451$). No other significant correlations are present although extraversion - NCA1 reaches the level $p = .058$ with $r = .319$ and neuroticism - NCN2 reaches the level $p = .052$ with $r = -.326$.

7.5 Discussion

In this chapter we examined the relationship between personality and game behaviour in the commercial video game Fallout 3. Since the goal of this research was to validate our results from Chapter 5, we compare the results of both investigations.

In Section 7.5.1 we compare the data-driven models. In Section 7.5.2 we compare the theory-driven models. In Subsection 7.5.3 we provide our interpretation of the results of the validation. In Subsection 7.5.4 general remarks are made concerning the results specific to this chapter. In Subsection 7.5.5 the control variables and their influence on the results are discussed.

7.5.1 Comparison of data-driven models

For our data-driven model of personality, our correlation and regression results in this chapter are higher than in Chapter 5. We can create linear models that explain over 75% of the variance for four of the five OCEAN traits. For the fifth trait (conscientiousness) we can explain almost 50% of the variance. This pattern is different for Chapter 5 where the extraversion is hardest to model. The effect sizes for both chapters are shown in Table 7.9.

Table 7.9 shows similar R^2 scores for the traits conscientiousness and agreeableness. Openness R^2 scores for both chapters are high, but Chapter 7 shows a somewhat higher

Table 7.6: Minimum and maximum correlations per group.

	Group 1		Group 2		Group 3		Group 4		Group 5	
Correlations	Pos.	Neg.	Pos.	Neg.	Pos.	Neg.	Pos.	Neg.	Pos.	Neg.
Openness										
Lowest							.415		.290	-.281
Highest							.415		.442	-.530
Conscientiousness										
Lowest								-.356	.315	-.303
Highest								-.356	.336	-.404
Extraversion										
Lowest							.319		.288	-.305
Highest							.319		.413	-.468
Agreeableness										
Lowest		-.432		-.420			.341		.280	-.281
Highest		-.432		-.482			.341		.394	-.438
Neuroticism										
Lowest		-.301				-.369		-.365	.467	-.282
Highest		-.379				-.369		-.410	.467	-.532

score. R^2 scores for Extraversion and Neuroticism are quite dissimilar.

Notable scores are openness and neuroticism for Chapter 7. Human behaviour is usually influenced by more than one variable and R^2 scores approaching the value 1.000 should be suspected of overfitting. Adding more participants to the sample may drastically change these scores because a large number of variables that have small influences in the model may be removed as error variance. So, the validation is debatable. The exact influence of the removed variables can only be interpreted in a new analysis.

7.5.2 Comparison of theory-driven models

When comparing the results of the Chapter 6 theory-driven analysis to the analysis in this chapter we may conclude that none of the correlations found in the Chapter 6 correlation analysis returns in this chapter. In Chapter 6 we see one strong correlation between agreeableness and NCA1 and correlations between neuroticism and NCC2, NCN1, and NCN3. In this chapter we see strong correlations between agreeableness and NCE2 and NCN3. The difference in findings leads to the conclusion that our investigation in this chapter does not validate the findings from Chapter 6.

Below, we present three factors that may be related to the difference in correlation effects: (1) the sample in this chapter is small relative to the effect sizes we expect to find for personality effects; a power analysis shows that there is a chance of 55% of missing effects that are present in the population when using 36 participants, (2) the game environment is different from the environment in the Chapter 6 game and may have led to different behaviours by our participants; a notable example of the differences is the amount of possible friendly conversation options; there was more opportunity to behave in a friendly manner in

Table 7.7: Descriptives of the theory-driven variables.

	N	Min	Max	M	SD
NCO2	36	1	8	3.64	1.807
NCC1	36	0	8	1.39	1.460
NCC2	36	10	19	13.08	2.419
NCE1	36	13	29	21.08	3.996
NCE2	36	2	33	8.53	6.500
NCA1	36	-6	10	2.22	3.642
NCN2	36	0	15	8.00	3.061
NCN3	36	1	25	7.14	5.431

Table 7.8: Correlations between theory-driven variables and personality traits.

	Openness	Conscientiousness	Extraversion	Agreeableness	Neuroticism
NCO2					
NCC1					
NCC2					
NCE1					
NCE2				-.444**	
NCA1					
NCN2					
NCN3				-.451**	

** $p < 0.01$, * $p < 0.05$ (not in the table)

the previous chapters' conversation; and (3) the variables have been built up in a different fashion; raw conversation variables have a maximum value of one instead of unlimited and movement is done using a combination of keyboard and mouse inputs instead of mainly mouse inputs as in the Chapter 6 game.

Our comparisons have taught us that, in order to create a game in which personality can be measured, two additions to the experiment are preferable. (1) Provide conversation options in which many of the traits may be expressed. In the current version, Fallout 3 presents a large number of unfriendly conversation options which lead to a limited number of expression traits. (2) Provide a large number of opportunities for players to move around in a non-linear fashion if they desire to do so. In Fallout 3, movement through the story was mainly linear which may have reduced the player's ability to explore freely.

7.5.3 Explaining the differences in effects

We present three possible reasons for the differences between the two studies: (1) the current experiment contained variables that were differently set up and fewer in number, (2) the contents of the games were rather different, and (3) the lower number of participants may cause an overestimation of the regression and correlation scores.

Table 7.9: Comparison between Chapter 5 and Chapter 7 R^2 scores.

Trait	R^2 Chapter 5	R^2 Chapter 7
Openness	.768	.982
Conscientiousness	.559	.486
Extraversion	.351	.837
Agreeableness	.724	.752
Neuroticism	.568	1.000

(Ad 1) The variables used in the current research and the variables used in the study of Chapter 5 are different in two ways: (a) in the Chapter-5 study a self-created game is used in which conversation variables are selected in such a way that they suited one of the personality traits; this is not the case in the current research, and (b) in the study of Chapter 5, movement variables are placed on locations where the player comes often and locations where the player does not come often; in the current research most triggers are placed at places where the players do not have to go to and thus do not come often.

(Ad 2) With regards to the difference in contents a clear example is that conversation in the NWN study is set-up to provide a variety in the possible answers, while the Fallout study contains conversations with some characters in which most of the answers are denials and sarcastic comments. The NWN study is mainly set in an external environment, while the Fallout study takes place entirely indoors.

(Ad 3) Pertaining to the overestimation of regression and correlation scores we may state that larger data samples improve analyses because random variations are evened out. In experiments with a large number of variables these are bound to be coincidental effects in error variance that disappear when more data is added. The larger effect sizes in the regression analysis may be due to overfitting because of the number of participants. The results in this research that overlap with the results in the research of Chapter 5 are most credible because they validate one-another.

In the case of the game contents, there are cosmetic and story differences between the NWN study and the FALLOUT study. In the FALLOUT study several story options, such as the possibility to warn multiple different NPCs, were absent. The setting in which the games took place is different as well. In NWN we see open environments, the outside, and a medieval village setting. In FALLOUT 3 we see inside areas, metal tunnels, and a futuristic, post-apocalyptic setting.

7.5.4 General remarks on personality traits

The point touched upon in Subsection 5.5.7 (i.e., similarities between conversations and multiple-choice questions) holds here as well. The 60 items in the questionnaire are descriptive statements on which the participant has to rate himself from one to five. In the FALLOUT study, conversation choices are responses to what other NPCs state. These statements cannot be compared with the descriptive statements within the personality questionnaires. The player's personality traits are shown by his in-game behaviour rather than by the player describing himself.

To our knowledge there have been no research investigations comparing the behavioural differences between a game and 'real' life. For example, when the player swears at an NPC, that does not necessarily mean that they would do such a thing in 'real' life. But it does mean that the player swears at an NPC now or that he does not think that it is a problem within a certain game situation. Even when someone behaves differently in a game, personality still has its influence on their in-game behaviour.

A challenging issue was how to prevent participants from choosing answers randomly instead of considering their choices carefully. In the study by Schreurs (2009) the participants received the instruction to behave in the game as they would behave in reality. In the current research we chose to follow a different perspective, because such an instruction is introspective, i.e., unrealistic (the situations within the game are not realistic) and not necessary (because personality is always involved). For the current research we explained in the instructions that the participants have to act in the same way as they would do when they were going to play the game till the end. In this way we tried to strengthen the ecological validity. Still, the participants knew that their choices are important because it would determine the further proceedings of the game.

7.5.5 Control variables

A number of correlations above .300 are found concerning the control variables. First, the significant correlations between the control variables themselves are discussed. The control variables provide insight into other possible factors besides personality that could have influenced the results.

Influences between control variables

Correlations can be found between the control variables themselves, as shown in Table 7.10. We observe that females have less experience with computers and computer games, and have more negative experiences with the controls of Fallout 3 and its game proceedings when compared to males. The argument may be made that the more experience one has with computers and computer games, the less difficult one might experience the controls. We note that participants who have much experience with computer games, liked the game more than others. These participants would not have that much experience with computer games, if they would not like computer games. There is a possibility that the easier one proceeds within the game, the more the game is enjoyed.

7.6 Chapter conclusions

The present chapter investigated the possibilities of using a virtual game to profile the personality traits of players. The answer to RQ 5: *To what extent can our models of personality in games be validated in different games?* is as follows. For our data-driven model of personality, our correlation and regression results in this chapter are higher than in Chapter 5. We can create linear models that explain over 75% of the variance for four of the five OCEAN traits. For the fifth trait (conscientiousness) we can explain almost 50% of the variance. This pattern is different for Chapter 5 where the extraversion is the hardest to

Table 7.10: Correlations between the control variables themselves.

	Age	Edu	Sex	Exp. comp.	Exp. games	English skills	Exp. ctrls	Exp. proc.	Fun
Age	x	.643							
Education	.643	x							
Sex			x	-.399	-.749		-.731	-.502	
Exp. computers			-.399	x	.519	.349	.516	.458	.333
Exp. games			-.749	.519	x	.417	.734	.629	.477
English skills				.349	.417	x			
Exp. controls			-.731	.516	.734		x	.722	
Exp. proceedings			-.502	.458	.629		.722	x	.368
Fun				.333	.477			.368	x

model. For all five traits, significant correlations have been found with in-game behaviour. More specifically, 87 of the total of 192 game variables showed significant correlations with personality traits.

The validation of our theory-driven variables was not successful. The effects found in this chapter do not overlap with the effects found in Chapter 6. For statistical reasons, our future approach should focus on (1) more experimental control in order to exclude interfering variables and (2) more participants in order to increase the power of the statistics used. Both the results for regression as well as for correlation are possibly influenced by the number of participants used in these investigations.

In the next chapter we will discuss heterodoxies and surprising effects found during the course of our research. We will also discuss our perspective on psychological phenomena in relation to games and gaming.

Chapter 8

Modelling in games

After seven chapters with a variety of research results, the idea to quantify individual player differences is still a challenging topic. To strengthen our grip on this topic we concentrate in this chapter on the role of models, in particular when we provide the models with some fixed structure so that they act as a framework for the researcher. In our case, we aim at models providing a framework by which we may attempt to interpret the psychology of players.

Definition 8.1 (Model) A model depicts how various behavioural characteristics, cognitive elements and emotional factors work in an interconnected way to generate an outcome with respect to behaviour, cognition, and emotion.

Definition 8.2 (Framework) A framework connects a set of ideas, principles, and rules in a harmonious manner to facilitate handling of situations. It will ensure that the boundaries are well matched so that desired results are achieved without any one of the characteristics, elements or factors overpowering or overshadowing others.

The framework with which we want to explain and predict player behaviour is composed of three parts: (1) the characteristics of players (in Section 8.1), (2) the responses of players (in Section 8.2), (3) the effects of a player's environment (in Section 8.3). Each of these parts can be described with one or more models.

In this chapter we further discuss: the psychometrics of psychology and games research (in Section 8.4), the adaptation of game content (in Section 8.5), and some pitfalls of applying psychological models in game research (in Section 8.6). We provide the chapter conclusions in Section 8.7.

8.1 Player characteristics

We identify two types of player characteristics: (1) stable traits and (2) semi-stable traits. Below, we first discuss the stable traits. Personality (8.1.1), intelligence (8.1.2), and demographics (8.1.3) are three traits that are considered to be stable over long periods of time, sometimes decades. Secondly, we discuss semi-stable traits. Skill (8.1.4) and preferences

(8.1.5) are considered to be semi-stable traits because they can be stable over longer periods of time but they can also change rapidly.

8.1.1 Personality

Personality has been extensively explained in Chapter 2. The investigations in Chapter 4 to Chapter 6 show that adapting games for personality assessment is a challenging affair.

When setting up the research for this thesis we expected personality to show effects in games similar to personality in the real world. When we designed the inn area in the scenario for Chapter 5, we expected extraverts to have higher movement scores in this area compared to introverts. This effect was not found. We had no previous expectations for conscientiousness regarding the completion of side stories. However, we did find a conscientiousness effect.

In retrospect, an explanation for both the abscence and the presence of these effects may be available, but a posteriori explanations are only characterised as good scientific practice for new theories and for designing and running new experiments. Moreover, we would like to be able to make adequate predictions in advance.

8.1.2 Intelligence

In the experiments of this thesis we did not test intelligence. However, we consider intelligence a candidate for future research because it ties into the development of new skills and the improvement of existing skills. Therefore we feel that intelligence is an important subject for consideration in games research. One way to measure intelligence is the Intelligence Quotient test (or IQ test, not to be confused with incongruity questionnaire).

If the intelligence of a player is known, it could be used as a scaling factor for the difficulty of game content. Intelligent players are likely to solve puzzles faster and they are also more likely to improve their skills faster. Suitable measures of intelligence in games may come in many forms. If a measure of the players skill in an early part of a game can be made comparable to a player's skill in a later part of the game then intelligence may be measurable. This comparison may need to be adjusted for the amount of time spent practicing the game and the time in between sessions. If a player in a home setting is playing other games besides the game in the experiment the improvement in skill may not accurately show his intelligence but rather a general practice effect.

8.1.3 Demographics

In Chapter 4 we saw the effects of the trait extraversion on game behaviour, in Chapter 5 we saw the effects of all five personality traits in a data-driven method, in Chapter 6 we saw the effects of the five traits in a theory-driven method, and finally, in Chapter 7 we saw the effects of the five factors in a different game. In parallel to measuring the personality traits in these chapters we also measured the demographic data of players (i.e., age, gender, etc.). We came to the conclusion that in many cases the demographic data had a significant impact on behaviour, in a few cases even a larger impact than the personality traits had.

Hence, we may conclude that collecting of demographic data may be effective and may even be essential in the unravelling of the causative factors of behaviour in games.

8.1.4 Skill

Skill in games improves with practice. An influential factor in the speed by which skill improves is intelligence. Measuring skill is relatively straightforward in games.

Skill can be represented by the combination of (1) speed of play, (2) number of errors, and (3) number of successes. Players with high speed, low error, and high success can be said to be skillful. When defining skill in this way, skill cannot be observed as an absolute value. When comparing a difficult task to an easy task, the absolute values of the measured skill will be different for each task. Therefore, skill can only be objectively assessed in relation to the skill of other players in combination with the task. An interpretation of skill can be found in Chapter 3 in the form of incongruity. The mental model of a player may be interpreted as his skill level.

8.1.5 Preferences

Preferences are formed by experiences and by adopting the preferences of others. Preferences may influence an experience by providing a positive or negative predisposition. Usually, humans are verbally able to express their preferences clearly. Inferring preferences from behaviour may prove difficult in situations where multiple variables provide an influence.

When attempting to measure preferences in games unmeasured variables may cause problems. To substantiate the idea of unmeasured variables, we provide the following example. In a game situation a player may prefer a forest route to a village route. Perceived danger on the forest route may lead the player to select the village alternative. In this way, observing the choice of route may incorrectly lead to the conclusion that a player prefers village routes. Here, the unmeasured variable "perceived danger" may cause an incorrect conclusion.

Researchers should be aware of the problem of unmeasured variables. If a researcher is interested in measuring preferences, consideration should be given to control variables that allow for the disentanglement of these effects.

8.2 Player responses

Actions and responses in games belong to the category of player properties that are extremely dynamic. For experiments, it is therefore important to develop a framework in which the outcome of behaviour, cognition, and emotion are fixed within boundaries. We consider a behaviour, cognition, or emotion as a short term event in the experiment, and so it may be possible to see what their value is within a defined range of possibilities.

8.2.1 Behaviour

Game behaviour is the basic measurement unit of game psychology. Since games usually do not present players with in-game questionnaires about their psychological state, the psychological state of the players is inferred from their behaviour in the game. Examples of game behaviour are key-presses and mouse movements. A slightly higher-level game behaviour could be the route that the player takes through a game area. Below we provide three definitions of these levels of behaviour (general, low-level, and high-level).

Definition 8.3 (General behaviour) General behaviour is defined as a measureable action performed by a human. Behaviours may range from simple (e.g., flexing a single muscle) to complex (e.g., creating a painting).

Definition 8.4 (Low-level behaviour) The low-level behaviour is a behaviour that can be recorded through the keyboard and mouse computer inputs (i.e., keypresses, mouse movement, and mouse clicks).

Definition 8.5 (High-level behaviour) The high-level behaviour is a series of low-level behaviours that are executed in pursuit of plans or goals.

Interpreting game behaviour is difficult and will be based on the topic of investigation. The question: "on average, how often does a player jump during world 1 of super mario brothers?" is easily answered but it does not provide much insight into a player's psychological state. The question: "does a player feel scared after playing world 1 of super mario brothers?" concerns an emotional state, but is much harder to answer when only looking at low-level game behaviours.

We consider it reasonable to classify the recorded averages of damage in the incongruity study as a high-level behaviour since the averages already represent a series of complex maneuvers to defeat an enemy while receiving as little damage as possible. In our experiments in Chapters 5 through 7, we investigated low-level behaviour by recording movement and conversation choices. Both types of low-level behaviour (movement and conversation) had the potential to measure high-level behaviour but we did not explicitly attempt to do so. However, our pooled variables were expected to record high-level behaviours. We expected our pooled variables to record preferences for locations, specific characters, and general preferences such as impatience.

In our experience, identifying which psychological construct is measured by a high-level game behaviour, as well as which high-level behaviours can be extracted from low-level behaviour, are two of the challenges of the experimental design process in games research. We expect that careful validation of game behaviour for a psychological construct is an essential task in game research that has not yet been properly addressed.

8.2.2 Cognition

Analogously to behaviour we start this subsection with a general definition of cognition. We do not subdivide cognition into lower levels of cognition since we only deal with cognition in general terms.

Definition 8.6 (Cognition) Cognition is defined as the process of human thought.

Cognitions are hard to investigate in general. This is also true for games. Cognitions may not show any kind of behavioural response, body language, or facial expression. We note that the most common way of investigating cognitions is by requesting a self-report from a person. One of the reasons cognitions are considered to be hard to investigate is that they are influenced by myriads of variables, both inside the person and in the situation (environmental). Cognitions may also be changed by the process of introspection or they may be inaccessable to introspection.

In games, there are many ways to observe the results of cognitions. We mention two of them. The first way is in the form of choices made in puzzles; the other way is in inspecting variations in plans. Often, the specific cognitions which a player has may not be relevant to game design or research. More often, the focus of game research lies in the resulting feelings or attitudes formed during play. Cognitions may be responsible for some of the choices that players make. Therefore, we may wish to know which cognitions are evoked by specific situations in games. Although it is a quite interesting area, we leave this topic for further research to future investigators.

8.2.3 Emotion

A comprehensive explanation of the emotional model used in this thesis has been explained in Subsection 2.3.1.

A prime goal in story-focussed games design is to provide players with an emotional experience. The option to know whether the player actually had an emotional experience would be a great benefit to game designers, particularly for cases when a game situation was designed to evoke emotion. There is currently not a comprehensive model of game behaviour that shows emotional effects. Some researchers have used additional tools such as cameras and physiological measurements in order to see if players show an emotional response (Kivikangas et al., 2010). The results were promising but still require additional research effort.

8.3 Player environment

Player environment can be described as the contents of the game that is played as well as to the environment in which the game is played. Content of the game refers to aspects such as the player's perspective (1st person, 3rd person, top-down), the mood of the game (colours, setting, music), the amount of freedom of choice given to the player, the mechanics used during play, and the design of levels and characters. By 'environment' we refer to factors such as: playing a game in an experimental setting or at home, the lighting conditions, outside noise during play, and the time at which a game is played. In this thesis we have not studied systematically the player's environment. We did study the impact of the use of colours in games on player emotions (Joosten et al., 2012), although we did not relate it to the incongruity or personality research in this thesis.

8.4 Game psychometrics

Psychometrics is the field of psychological assessment. Psychometrics investigates the properties of psychological tests.

We discuss the following topics: the possible spectrum of behaviour in games (8.4.1), the relationship between low-level actions and high-level implications (8.4.2), experimental control in game research (8.4.3), the risks in implementing multiple game situations in one experiment (8.4.4), the re-use capacity of game-based metrics (8.4.5), the relationship between actual behaviour and game behaviour (8.4.6), the scope of the given predictions as investigated by the game models (8.4.7), validation of psychological findings in game research (8.4.8), and the use of variables to improve experimental control (8.4.9).

8.4.1 Spectrum of behaviour

Players have a limited spectrum of behaviour to display. As stated before, a player can only display his behaviour to a computer by using the input peripherals. Behaviour can vary in frequency, speed, and time (to name only a few). However, some involuntary responses, such as those caused by changes in the player's physiology, may not be directly available for display.

Changes in response frequency may often be related to situations on the screen. The appearance of an adversary is likely to increase the frequency of attack-button pressing. Differences in response frequency between players may show some difference in player psychological state, such as skill or reaction speed.

8.4.2 Low-level actions, high-level implications

Low-level behaviour can, when coupled with additional information, be useful to investigate high-level psychological processes. Movement toward or away from different visible stimuli can provide information about the player's plans or about his preferences. For instance, the speed of movement through dark areas may be informative on the player's uncertainty or anxiety. Our research in Chapter 5 shows that the amount of movement in a colourful and strange area of the game is illustrative for the extraversion of the player.

A player's behaviour in story-related situations may hold relevant high-level information. If two different scenario options are offered to players, the choice of a scenario may display a player's preferences or expectations. It may also show personality traits related to fear or curiosity.

8.4.3 Experimental control

In an experiment where players play a complex game, it may be virtually impossible to control for all the variation in behaviour unless (1) the experiment has an extremely large number of participants, (2) the game has effective control variables, or (3) the psychological construct that is studied has a very significant (in the order of $p < 0.01$ or smaller) correlation to game behaviour. Experiments in which an attempt is made to fit a curve to a large number of variables run the risk of 'overfitting' the data (see Subsection 7.5.1).

8.4.4 Multiple different game situations

When an experiment is performed by using a game, it may be tempting to implement multiple different situations in order (1) to allow a player freedom of choice or (2) to investigate a wide variety of behaviours. A consequence of using multiple situations is that a large number of subjects is needed to gather data in all situations. Potentially, each situation in a game could influence the other situations in the game.

A more practical approach would be to design a single game situation and to measure a single variable. However, such an approach may reduce the ecological validity of a study by making the experiment quite different from the game in which the techniques may eventually be used. An alternative is to test a large number of subjects. This could be done by offering the game in an online venue. The disadvantage of offering the game in an online venue is that it is harder to perform additional tests such as personality questionnaires.

8.4.5 Replay value and repeated measurements

When considering a game as a psychological metric that can be applied multiple times to the same participant, it may be important to consider the type of game situation that has been chosen. In story-type games, especially highly intelligent players may have an exact recollection of the game progress. When a player arrives in a situation that is exactly the same as a situation encountered before, his behaviour may be different (owing to learning and adaptation) and therefore, the results of the second experiment may need to be re-evaluated with respect to being a psychologically reliable measure. If a game is selected in which a player must solve puzzles as fast as possible, seeing the same puzzles for a second time will most likely make the player faster. His entire solving behaviour may also be much more effective because he already knows the solution. In games of skill, if a player has been playing other games between a first and a second playthrough, the player may have increased his skill because of those other games or he may even have changed his playing style. Careful consideration should go into the reasons for building tests into games as well as careful validation that takes into account the natural development of players before attempting to create a game-based test that will be used multiple times.

8.4.6 Comparison to actual behaviour

Placing a person in a park inside a game may evoke a quite different behaviour compared to putting that same person in a real life park. So far, we have not attempted to compare the behaviour in game situations to similar situations in real life. There is literature suggesting that behaviour in virtual environments may be similar to behaviour in real life situations (see Kozlov and Johansen, 2010) but this investigation has not been carried out by comparing video game situations to real life situations.

8.4.7 Individual effects versus mass effects

The results obtained in this thesis are based on statistical methods that investigate effects on a group statistical level. The correlation effects found allow us to draw conclusions about the influence that personality has on game behaviours. The predictions by us here refer to groups of people rather than to an individuals' specific choices.

In future research, we would like to investigate what facets of personality interact with facets of a game to evoke behaviour. We consider behaviour in a game to be the result of a complex interplay of facets that need to be unravelled before it may lead to effective and personalised adaptations to games. An example is that extraversion is based on six facets. A score of extraversion is therefore a compounded product of these facets which may lose the specificity needed for adaptation to an individual.

8.4.8 Validity

Validity is an important and recurring theme in this thesis. In Chapters 4 through 7 we investigate possible applications of games as personality tests. In order for a game to be a successful personality test, the game needs to be validated. By 'validated' we mean that the game needs to be able to measure personality as it functions in 'real' life.

In order to perform validations in this thesis, we have investigated the correlations between game variables and the traits measured by the NEO-PI-R personality questionnaire in three different games (Chapters 4, 5, and 7) and in three different approaches within games (Chapters 4, 5, 6, and 7). These validations have produced some successful results. However, the most thorough form of validation is to show the relationship between personality-related behaviour in a game and personality related behaviour in real life directly. If a person is shy in real life but not in a game the validity of a game in the function of a personality test becomes questionable for that part of personality. The focus on validation using real-life personality is required in future research if games are to be used as true personality tests.

8.4.9 Control variables

In our research for the Chapters 4 to 7, we have conducted investigations into the demographics of our test samples in parallel to our primary investigation.

Our experiments were set up in such a way that personality was expected to produce the largest results. Our research shows that the demographics produce large correlation effects on game variables. In future work, control variables should already be taken into account at the design phase of the experiment. This may result in the necessity for much larger sample sizes.

8.5 Application of personality theory in games

Below, we discuss two possible applications for personality in games: (1) games as personality tests and (2) personality-based game adaptation.

8.5.1 Games as personality tests

With further investigation, games may be suitable for use as personality tests. Games may alleviate some of the drawbacks of contemporary personality tests. Using games, large quantities of behavioural data may be collected cheaply and quickly. Games are also rather suitable for testing children because of the medium.

Three obstacles when using games are: (1) the long development cycle required to validate the games if a large number of datapoints are used, (2) games also require expert knowledge-based supervision of their content in order to see if all personality traits are expressible, and (3) the interpretation of the meaning of different game situations in terms of personality.

8.5.2 Personality-based game adaptation

If a game is filled with content that coincides with a player's preference (compared to a generic game) then the preference-filled game will be rated more positively. Following this line of reasoning we could consider games that can adapt their content to a player's preferences as desirable games. In our research in the Chapters 4 through 7 we see that personality leads players to choose some specific options in the game over others. We believe that we may state that these patterns of choice indicate what a player's preferences are.

Moreover, it may be possible to develop a game with lists of potential situations and choices for each personality configuration. The choices can be provided to the player as soon as a player's personality is indicated by his behaviour in the game. Dynamically altering a game in this way should increase the game's attractiveness to each player for which it can provide adequate content for his personality. Several techniques have been developed for adapting games. An example of such a technique is dynamic scripting Spronck et al. (2006). Matching a technique to the task of personality-based adaptation will require further consideration.

An investigation by Schreurs (2011) indicates that adaptation based on personality is challenging. In Schreurs' investigation, the participant's personality was measured and correlated to the reported appreciation of three in-game situations (a fight, a puzzle, and a social challenge). The participant's demographics were also measured. The investigation shows a few correlations between personality and situation appreciation. However, the investigation shows that demographic factors such as skill and gameplay experience correlate significantly in multiple situation-personality combinations (for the full details please consult the work by Schreurs (2011)). Therefore, demographics should be considered as control variables. However, we maintain that personality-based game adaptation shows much promise and that this line of research deserves attention in the future.

8.6 Pitfalls in applying psychological theory to games

When psychological processes are investigated in computer science the lack of validation of an applied measure is a common oversight (Andrade et al., 2006). Experiments claiming to influence any psychological process, such as an emotion, should always be thoroughly validated. One possible means of validation is by showing an actual change in emotion in human participants when applying the measure.

An example of the validation of a measure in this thesis is the incongruity research in Chapter 3. To our knowledge, the emotional responses predicted by the incongruity theory had not been validated for games at the time at which we conducted our experiments.

Prior to the start of our experiments we considered our experimental measure to accommodate the three states of incongruity (high positive, high negative, and low incongruity) properly. Because of this assumption, we expected to see the emotions boredom, frustration, and pleasure to be clearly represented in our data. However, when we looked at the results, we could see that only one of the three emotions showed the expected pattern for the three incongruity states, viz., frustration. We propose two possible reasons for the lack of a significant difference in boredom between our experimental conditions: (1) our experiment was not sufficiently easy, or (2) the incongruity theory is not predicting the emotion correctly. For pleasure, similar reasons hold: (1) our game might be badly designed, or (2) the incongruity theory might not be predicting the pleasure emotion correctly. In other words, our experiment shows that expectations of the workings of a theory are not necessarily warranted.

While investigating incongruity, we found some confirmation for the psychological concepts. In comparison, in Chapter 4 we implemented extraversion concepts found in the literature. These concepts functioned differently from what was expected based on the literature. In some cases the difference was the exact opposite of what was expected.

One lesson that can be learned from these investigations is that we need to operationalise every detail of a theory thoroughly. For instance, incongruity theory may state that differences in complexity cause emotional responses but it does not define clearly what the complexity is per topic, how they differ per topic, and if there is a basic complexity which holds for all emotions. In summary, we have no clear view of what the level of complexity is for a game with an entirely empty world with just a finish line at one end. We might suppose that the control of the player in the game carries some level of complexity, but at this time there is no quantifiable theoretical definition to work with. For example, we assumed that our game algorithm would balance itself, but we did not sufficiently account for the possibility that even one single enemy might still be too difficult for some (or many) of the players.

All in all, in future research more attention should be paid to validate to what extent the intended subject is actually measured. When computer science research attempts to handle psychological problems, a common mistake is to assume that descriptions in the literature function in the same manner in games as they do in the real life. We advise caution when implementing psychological theories in novel approaches. Researchers should extensively test and pilot their experiments before starting a full experiment

8.7 Chapter conclusions

In this chapter we discussed modelling in games. In particular we investigated the idea of developing a framework and subsequently testing its capabilities. From the discussion, we may conclude that game research provides many challenges for technological development but also many great opportunities for investigating human psychology.

In the next chapter we will provide the concluding remarks to this thesis.

Chapter 9

Conclusions

In this chapter we provide an answer to the research questions and the problem statement proposed in Chapter 1. We reiterate the research questions (9.1) and provide their answers. Then we answer the problem statement (9.2). Finally we discuss future research (9.3).

9.1 Answering the research questions

In Section 1.2 we formulated five research questions. This section provides answers to those questions.

> **Research question 1:** *To what extent are games suitable for measuring incongruity?*

The answer to the first research question is derived from Chapter 3. Incongruity theory states that emotions are felt when incongruity arises. We are able to find one of the three emotional reactions (viz. frustration) in relation to the incongruity level and one partially consistent emotional reaction (viz. pleasure). For the reactions that do not arise as predicted, we found that frustration increases with difficulty, while the pleasure value remains roughly the same for easy and balanced difficulty, but drops drastically for hard difficulty. We found that boredom does not conform to the predictions of the incongruity theory. We propose that for our game the easy condition might be too hard in order to provoke the predicted boredom reaction. Alternatively, it may also be possible that subjects find the easy condition boring but that they display coping behaviour (like looking around or daydreaming) which mediates boredom. Our answer to the research question is that in our game we may conclude that players feel frustrated when playing a hard game and that they feel pleasure when playing a balanced game. We are unable to conclude that players feel boredom when playing an easy game.

> **Research question 2:** *To what extent can games be used to measure complex psychological traits such as extraversion?*

The answer to the second research question is derived from Chapter 4. Investigating the question of correlation between in-game behaviour and personality scores on the NEO-PI-R, a test was administered to a pool of 39 participants and yielded outcomes for the 21 in-game elements. The outcomes were analysed for correlations using regression analysis. Our answer to the research question is that to a large extent we may conclude that it is possible to measure complex psychological processes such as extraversion and its facets, using behaviour in a virtual world.

> **Research question 3:** *To what extent can a data-driven personality profile be created based on game behaviour?*

The answer to the third research question is derived from Chapter 5. Based on our investigation we may conclude that personality effects on game behaviour exist for all five traits of the FFM. We found these effects when we performed linear regression analysis and when we performed detailed correlation analysis of the game behaviour variables to scores on a personality questionnaire. Therefore we may conclude that we are able to produce estimates of a participant's personality based on the game variables. Of our 260 game variables, 57 variables proved to correlate with one or more personality traits. Our answer to the research question is that to a modest extent we are able to create a personality profile based on game behaviour.

> **Research question 4:** *To what extent does a theory-driven model explain personality in games?*

The answer to the fourth research question is derived from Chapter 6. Based on our tests with the theory-driven model of personality we may conclude the following. The tests of the theory-driven model reveal correlations between model variables and personality traits for four of the eleven model variables. The correlations found numbered three for the neuroticism trait and one for the agreeableness trait. We showed that statistical relationships exist between our theory-driven variabeles and the five personality traits and we also showed that the percentage of correlations for the theory-driven model surpasses the percentage of correlations found for the data-driven model. Our answer to the research question is that we can only in part explain personality in games using a theory-driven model.

> **Research question 5:** *To what extent can our models of personality in games be validated in different games?*

The answer to the fifth research question is derived from Chapter 7. From our investigations we may draw two conclusions.

Firstly, for our data-driven model of personality, we can create linear models that explain over 75% of the variance for four of the five OCEAN traits. For the fifth trait (conscientiousness) we can explain around 50% of the variance. This pattern is different for the game of Chapter 5 where the extraversion is the hardest to model. For all five traits, significant correlations have been found with in-game behaviour. More specifically, 87 of the total of 192 game variables showed significant correlations with personality traits.

Secondly, the validation of our theory-driven variables was not successful. The effects found in this chapter do not overlap with the effects found in Chapter 6. Our answer to the research question is that, although the statistical results for the investigations show personality effects in different games, additional research is required to validate personality in games.

9.2 The problem statement

In this section we answer the problem statement of this thesis which was formulated in Section 1.2. We base our answer on the conclusions of the research questions provided in the previous section.

> **Problem statement:** *To what extent are games an appropriate means for measuring the differences between individuals based on psychological theories?*

In this thesis we have investigated two different psychological models, the incongruity model and the personality model. In order to investigate the personality model we have used five different approaches: (1) the explicit use of a game as a digital testing environment, (2) the use of a data-driven approach in a game scenario that resembles commonly found game scenarios in commercial games, (3) a theory-driven approach in the same game as used in Chapter 5, and both (4) a data-driven as well as (5) a theory-driven approach in a commercially available and popular game. Our approaches on the two psychological models have yielded results with varying degrees of success.

From all the tests and experiments, we may conclude that all methods used would benefit from a larger test sample. We also may conclude that collecting such a test sample using our methods will require a long validation process. However, we showed that each of the methods provides sufficient results to consider the research a worthwhile and scientific step forwards. We feel that this research provides many leads for further research as well as for revealing some of the pitfalls that researchers should be mindful of when they consider investigations into psychological models in games.

Our answer to the problem statement is that we consider games to be an appropriate means for investigating individual differences based on psychological theories given the fact that we have found both the theory-driven as well as the data-driven correlation results and regression results for personality in games. We have also found relationships between emotions (frustration and pleasure) and game difficulty as predicted by incongruity theory.

9.3 Future research

We have provided several venues for future research in Chapter 8. Here, we provide three suggestions for future work along with four requirements for future research.

1. Is personality expressed in the same manner in games and in real life? Our research indicates that there may be differences in the expression of personality. Innovative research may also be developed for new personality traits that are only expressed in video games.

2. The effects of demographics and other factors in games deserve attention in future work. In this thesis, indications are present that relationships between (1) demographics and (2) game preferences, and (3) game behaviour may be very strong.

3. The question of how to gather data in games that is of high quality and that consists of data from a large number of participants remains open. This thesis shows that there is a need for improvements in both quality and quantity of gathered data.

The most important requirement to make for future work is that (1) a large increase in the number of experimental participants is needed if a large number of variables is measured. If a correlation analysis is chosen, for each of the correlations tested one in 20 is incorrectly assumed to be significant at $p < 0.05$ so the significance threshold should be drastically lowered in order to be able to test a large number of correlations. Furthermore, (2) validation of results in different games and in different in-game situations are desireable.

For the purpose of creating player models in games, especially if the goal of such models is the improvement of entertainment in games, (3) control variables deserve attention. In the control variables we have seen that there are (unexpected) correlations to be found. This means that these variables need to be controlled for interaction with the primary experimental variables.

The final requirement for future reseach is that (4) results are validated in multiple games before they are considered to be generalisable. This thesis shows that validation may be difficult.

References

Allport, G. W. and Odbert, H. S. (1936). Trait-names: A psycho-lexical study. *Psychological monographs*, 47(1):i–171.

Andrade, G., Ramalho, G., Gomes, A. S., and Corruble, V. (2006). Dynamic game balancing: an evaluation of user satisfaction. In Laird, J. and Schaeffer, J., editors, *Proceedings of the Second Artificial Intelligence and Interactive Digital Entertainment Conference*.

Aponte, M.-V., Levieux, G., and Natkin, S. (2011). Measuring the level of difficulty in single player video games. *Entertainment Computing*, 2(4):205–213.

Arney, F. M. (2004). *A comparison of direct observation and self-report measures of parenting behaviour*. PhD thesis, University of Adelaide - Department of Psychiatry.

Baba, N. and Handa, H. (2007). Commons game made more exciting by an intelligent utilization of the two evolutionary algorithms. *Studies in computational intelligence*, 71:1–16.

Back, M. D. and Egloff, B. (2009). Yes we can! A plea for direct behavioural observation in personality research. *European Journal of Personality*, 23:403–405.

Back, M. D., Schmukle, S. C., and Egloff, B. (2009). Predicting actual behaviour from the explicit and implicit self-concept of personality. *Journal of personality and social psychology*, 97(3):533–48.

Bakkes, S. C. J., Spronck, P. H. M., and van Lankveld, G. (2012). Player behavioural modelling for video games. *Entertainment Computing*, 3(3):71–79.

Bartle, R. (1996). Hearts, clubs, diamonds, spades: Players who suit muds. *Journal of MUD research*.

Beume, N., Hein, T., Naujoks, B., Piatkowski, N., Preuss, M., and Wessing, S. (2008). Intelligent anti-grouping in real-time strategy games. In *CIG'08. 2008 IEEE Symposium On Computational Intelligence and Games*, pages 63–70. IEEE.

Billings, D. (2006). *Algorithms and assessment in computer poker*. PhD thesis, University of Alberta.

Billings, D., Papp, D., Schaeffer, J., and Szafron, D. (1998). Opponent modeling in poker. In *AAAI-98. Proceedings of the National Conference on Artificial Intelligence*, pages 493–499. John Wiley & Sons Ltd.

Bradley, M. M. and Lang, P. J. (1994). Measuring emotion: the self-assessment manikin and the semantic differential. *Journal of behavior therapy and experimental psychiatry*, 25(1):49–59.

Brügmann, B. (1993). Monte carlo go. Technical report, Max-Planck-Institute of Physics, Föhringer Ring 6, München, Germany.

Butler, S. and Demiris, Y. (2010). Using a cognitive architecture for opponent target prediction. In Romano, D. M. and Moffat, D. C., editors, *Third International Symposium on AI Games*, pages 55–62.

Campbell, M. (1999). Knowledge discovery in Deep Blue. *Communications of the ACM*, 42(11):65–67.

Canossa, A. (2008). Towards a theory of player: Designing for experience. Technical report, IO Interactive/Denmark's School of Design, Copenhagen.

Carmel, D. and Markovitch, S. (1993). *Learning models of opponent's strategy in game playing*. Technion-Israel Institute of Technology. Computer Science Department.

Chaslot, G., Bakkes, S., Szita, I., and Spronck, P. (2008). Monte-carlo tree search: A new framework for game ai. *Proceedings of the Fourth Artificial Intelligence and Interactive Digital Entertainment Conference*, pages 216–217.

Chin, D. N. (1993). Acquiring user models. *Artificial Intelligence Review*, 7(3-4):185–197.

Chou, T.-J. and Ting, C.-C. (2003). The role of flow experience in cyber-game addiction. *CyberPsychology & Behavior*, 6(6):663–675.

Cohen, J. (1988). *Statistical power analysis for the behavioural sciences*. Lawrence Erlbaum, 365 Broadway, Hillsdale, N.J., USA, second edition.

Cohen, J. (1992). A power primer. *Quantitative Methods in Psychology*, 112(1):155–159.

Costa, P. T. and McCrae, R. R. (1992). *Revised neo personality inventory (neo pi-r) and neo five-factor inventory (NEO-FFI)*. Psychological Assessment Resources, Odessa, Fla. PO Box 998, Odessa 33556.

Costa, P. T. and McCrae, R. R. (1995). Domains and facets: hierarchical personality assessment using the revised NEO personality inventory. *Journal of personality assessment*, 64(1):21–50.

Costa, P. T. and McCrae, R. R. (2008). *The SAGE handbook of personality theory and assessment*, chapter 5, pages 159–181. The Guilford Press, 72 Spring Street, New York, N.Y., USA.

Cowley, B., Charles, D., Black, M., and Hickey, R. (2008). Toward an understanding of flow in video games. *Computers in Entertainment*, 6(2):1–27.

Cowley, B., Charles, D., Black, M., and Hickey, R. (2009). Analyzing player behavior in pacman using feature-driven decision theoretic predictive modeling. In *CIG 2009. IEEE Symposium on Computational Intelligence and Games*, pages 170–177. IEEE.

Csikszentmihalyi, M. (1988). *The flow experience and its significance for human psychology.* Cambridge University Press.

Csikszentmihalyi, M. and Csikszentmihalyi, I. S., editors (1989). *Optimal Experience: Psychological studies of flow in consciousness.* Cambridge University Press, The Edinburgh building, Cambridge CB2 2RU, UK.

Csikszentmihalyi, M. and Hunter, J. (2003). Happiness in everyday life: The uses of experience sampling. *Journal of Happiness Studies*, 4(2):185–199.

Digman, J. M. (1990). Personality Structure: Emergence of the Five-Factor Model. *Annual Review of Psychology*, 41(1):417–440.

Donkers, H. H. L. M., Uiterwijk, J. W. H. M., and van den Herik, H. J. (2003). Admissibility in opponent-model search. *Information Sciences*, 154(3–4):119–140.

Drachen, A., Canossa, A., and Yannakakis, G. N. (2009). Player modeling using self-organization in tomb raider: underworld. In *Computational Intelligence and Games, 2009. CIG 2009. IEEE Symposium on*, pages 1–8. IEEE.

El-Nasr, M. S. (2007). Interaction, Narrative, and Drama - Creating an Adaptive Interactive Narrative using Performance Arts Theories. *Interaction Studies*, 8(2):209–240.

Elo, A. E. and Sloan, S. (2008). *The Rating of Chess Players, Past and Present.* Ishi Press.

Field, A. (2009). *Discovering Statistics Using SPSS (Introducing Statistical Method).* Sage Publications Ltd, third edition.

Fischer, G. (1999). User modeling: The long and winding road. In *Proc. User Modeling*, pages 349–356. Springer-Verlag.

Fisher, R. J. (1993). Social desirability bias and the validity of indirect questioning. *Journal of Consumer Research*, 20(2):303–315.

Fürnkranz, J. (2007). Recent advances in machine learning and game playing. *ÖGAI Journal*, 26(2):1–10.

Furr, R. M. (2009). Personality psychology as a truly behavioural science. *European Journal of Personality*, 23(5):369–401.

Geen, R. G. (1984). Preferred stimulation levels in introverts and extroverts: Effects on arousal and performance. *Journal of Personality and Social Psychology*, 46(6):1303–1312.

Glover, E. (1991). *Freud or Jung.* Northwestern University Press, Evanston, Illinois 60201, USA.

Goldberg, L. R. (1993). *The structure of personality traits: Vertical and horizontal aspects*, pages 169–188. American Psychological Association, Washington, DC, USA.

Goodman, R. (1997). The strengths and difficulties questionnaire: a research note. *Journal of child psychology and psychiatry*, 38(5):581–586.

Gosling, S. D., John, O. P., Craik, K. H., and Robins, R. W. (1998). Do people know how they behave? Self-reported act frequencies compared with on-line codings by observers. *Journal of personality and social psychology*, 74(5):1337–1349.

Gross, S. J. and Niman, C. M. (1975). Attitude-behaviour consistency: A review. *Public Opinion Quarterly*, 3(39):358–368.

Grüsser, S. M., Thalemann, R., and Griffiths, M. D. (2007). Excessive computer game playing: evidence for addiction and aggression? *CyberPsychology & Behavior*, 10(2):290–292.

Halim, Z., Fazal-Ur-Rehman, M., and Baig, A. R. (2010). Evolution of entertainment in computer games. *International Journal of Computer Applications*, 3(9):38–45.

Hawkins, D. W. (2004). The problem of overfitting. *Journal of Chemical Information and Computer Sciences*, 35(19):1–12.

Hoekstra, H. A., Ormel, H. J., and de Fruyt, F. (2007). *NEO-PI-R / NEO-FFI Handleiding.* Hogrefe Uitgevers B.V., Lutmastraat 101, 1073 GS Amsterdam, The Netherlands.

Hsu, C.-L. and Lu, H.-P. (2004). Why do people play on-line games? an extended TAM with social influences and flow experience. *Information & Management*, 41(7):853–868.

Hunicke, R. and Chapman, V. (2004). AI for dynamic difficulty adjustment in games. In *Proceedings of the Challenges in Game AI Workshop, Nineteenth National Conference on Artificial Intelligence*, pages 91–96.

Hunt, J. M. (1963). Motivation inherent in information processing and action. In Harvey, O. J., editor, *Motivation and social interaction*, pages 35–94, New York, NY, USA. Ronald Press.

Iida, H., Kotani, Y., and Uiterwijk, J. (1995a). Tutoring strategies in game-tree search. In *Games of No Chance*, pages 433–435. Cambridge University Press.

Iida, H., Nakagawa, T., Spoerer, K., and Sone, S. (2012). Three elemental game progress patterns. *Intelligent Science and Intelligent Data Engineering*, pages 571–581.

Iida, H., Takahara, K., Nagashima, J., Kajihara, Y., and Hashimoto, T. (1995b). An application of game-refinement theory to mah jong. *Entertainment Computing*, pages 253–282.

Iida, H., Uiterwijk, J., van den Herik, H. J., and Herschberg, I. S. (1993a). Potential applications of opponent-model search. part 1, the domain of applicability. *ICCA Journal*, 16(4):201–208.

Iida, H., Uiterwijk, J., van den Herik, H. J., and Herschberg, I. S. (1993b). Potential applications of opponent-model search. part 2, risks and strategies. *ICCA Journal*, 17(1):10–14.

John, O. P. and Srivastava, S. (1999). *The Big-Five trait taxonomy: History, measurement, and theoretical perspectives*, chapter 4, pages 102–138. Oxford University Press.

Joosten, E., van Lankveld, G., and Spronck, P. H. M. (2010). Colors and emotions in video games. In Ayesh, A., editor, *11th International Conference on Intelligent Games and Simulation GAME-ON 2010*, pages 61–65. EUROSIS.

Joosten, E., van Lankveld, G., and Spronck, P. H. M. (2012). Influencing player emotions using colors. *accepted for publication in the Journal of Intelligent Computing.*

Kivikangas, M., Ekman, I., Chanel, G., Järvelä, S., Cowley, B., Salminen, M., Henttonen, P., and Ravaja, N. (2010). Review on psychophysiological methods in game research. In Petri, L., Mette, T. A., Harko, V., and Annika, W., editors, *Proceedings of DiGRA Nordic 2010: Experiencing Games: Games, Play, and Players*, Stockholm. University of Stockholm.

Kline, P. (1999). *The handbook of psychological testing.* Routledge, 11 New Fetter Lane, London, UK, second edition.

Kobsa, A. (1993). User modeling: Recent work, prospects and hazards. *Human Factors in Information Technology*, 10:111–125.

Koch, N. and Rossi, G. (2002). Patterns for adaptive web applications. In *Proceedings 7th European Conference on Pattern Languages of Programs*, pages 179–194, Irsee, Germany.

Kolar, D. W., Funder, D. C., and Colvin, C. R. (1996). Comparing the accuracy of personality judgments by the self and knowledgeable others. *Journal of personality*, 64(2):311–337.

Koster, R. (2004). *A Theory Of Fun In Game Design.* Paraglyph Press.

Kozlov, M. D. and Johansen, M. K. (2010). Real behavior in virtual environments: Psychology experimentss in a simple virtual-reality paradigm using video games. *Cyberpsychology, Behavior, and Social Networking*, 13(6).

Laramee, F. (2002). Using n-gram statistical models to predict player behavior. In Rabin, S., editor, *AI Game Programming Wisdom*, volume 1, pages 797–804. Charles River Media, Hingham, MA.

Ledezma, A., Aler, R., Sanchis, A., and Borrajo, D. (2005). Predicting opponent actions by observation. *RoboCup 2004: Robot Soccer World Cup VIII*, pages 286–296.

Ledezma, A., Aler, R., Sanchis, A., and Borrajo, D. (2009). Ombo: An opponent modeling approach. *AI Communications*, 22(1):21–35.

Likert, R. (1932). A technique for the measurement of attitudes. *Archives of psychology*.

Lockett, A., Chen, C., and Miikkulainen, R. (2007). Evolving explicit opponent models in game playing. In *Proceedings of the 9th annual conference on Genetic and evolutionary computation*, pages 2106–2113. ACM.

Magnavita, J. (2002). *Theories of personality: Contemporary approaches to the science of personality*. John Wiley and Sons.

Markovitch, S. and Reger, R. (2005). Learning and Exploiting Relative Weaknesses of Opponent Agents. *Autonomous Agents and Multi-Agent Systems*, 10(2):103–130.

Marsland, T. (1986). A review of game-tree pruning. *ICCA Journal*, 9(1):3–19.

McCrae, R. R. and Costa, P. T. (1997). Personality trait structure as a human universal. *The American psychologist*, 52(5):509–16.

McCrae, R. R. and Costa, P. T. (2003). *Personality in adulthood: A five-factor theory perspective*. The Guilford Press.

Miller, D. (1977). Roles of naturalistic observation in comparative psychology. *American Psychologist*, 32(3):211.

Millington, I. (2006). *Artificial Intelligence for Games*. Morgan Kaufmann Publishers, San Francisco, California, USA.

Mischel, W. (1972). Direct versus indirect personality assessment: evidence and implications. *Journal of consulting and clinical psychology*, 38(3):319–24.

Mommersteeg, F. (2002). Pattern recognition with sequential prediction. In Rabin, S., editor, *AI Game Programming Wisdom*, volume 1, pages 586–595. Charles River Media, Hingham, MA.

Nakatsu, R., Rauterberg, M., and Vorderer, P. (2005). A new framework for entertainment computing: from passive to active experience. *Entertainment Computing-ICEC 2005*, pages 1–12.

Newell, A. and Simon, H. (1972). *Human problem solving*. Prentice-Hall Englewood Cliffs, NJ.

Pennebaker, J. W., Chung, C. K., Ireland, M., Gonzales, A., and Booth, R. J. (2007). *The Development and Psychometric Properties of LIWC2007*. LIWC.net.

Prensky, M. (2003). Digital game-based learning. *Computers in Entertainment (CIE)*, 1(1):21–21.

Rabin, S., editor (2008). *AI Game Programming Wisdom 4*. Charles River Media, Inc., Hingham, Massachusetts, USA.

Rauterberg, M. (1994). About the relationship between incongruity, complexity and information: design implications for man-machine systems. *Design*, pages 122–132.

Rauterberg, M. (1995). About a framework for information and information processing of learning systems. *Information System Concepts–Towards a consolidation of views*, pages 54–69.

Rauterberg, M. (1998). About a method to measure the ergonomic quality of user interfaces in a task independent way. In *Workshop on Emerging Technologies in Human Engineering, Testing and Evaluation. NATO Research Study Group*, volume 24.

Russell, J. (1980). A circumplex model of affect. *Journal of personality and social psychology*, 39(6):1161.

Ryman, D., Biersner, R., and La Rocco, J. (1974). Reliabilities and validities of the mood questionnaire. *Psychological Reports*.

Schreurs, S. (2009). Psychological profiling of game players. Bachelor's thesis, Tilburg University.

Schreurs, S. (2011). Personality and game preferences. Master's thesis, Tilburg University.

Schull, N. (2002). *Escape mechanism: Women, caretaking, and compulsive machine gambling*. Center for Working Families, University of California, Berkeley Berkeley.

Shannon, C. (1950). Programming a computer for playing chess. *Philosophical magazine*, 41(314):256–275.

Smith, R. (1977). Empirical science and value assumptions: Lessons from C. G. Jung. *Journal of Religion & Health*, 16(2):102–109.

Smith, R. (1981). Boredom: A review. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 23(3):329–340.

Spronck, P. H. M., Balemans, I., and van Lankveld, G. (2012). Player profiling with fallout 3. In *Accepted for the AIIDE 2012 conference*.

Spronck, P. H. M., Ponsen, M., Sprinkhuizen-Kuyper, I., and Postma, E. (2006). Adaptive game ai with dynamic scripting. *Machine Learning*, 63(3):217–248.

Spronck, P. H. M., Sprinkhuizen-Kuyper, I., and Postma, E. O. (2004). Difficulty scaling of game ai. In *Proceedings of the 5th International Conference on Intelligent Games and Simulation (GAME-ON 2004)*, pages 33–37.

Stege, L., van Lankveld, G., and Spronck, P. H. M. (2011). Teaching electrical engineering with a serious game. In Wiemeyer, J. and Göbel, S., editors, *Theory, Technology & Practice: Proceedings of the GameDays 2011*, pages 29–39.

Stege, L., van Lankveld, G., and Spronck, P. H. M. (2012). Teaching high school physics with a serious game. *International Journal of Computer Science in Sports*, 11(1).

Susi, T., Johannesson, M., and Backlund, P. (2007). Serious games–an overview. *Skövde: University of Skövde (Technical Report HS-IKI-TR-07-001)*.

Sweetser, P. and Wyeth, P. (2005). Gameflow: a model for evaluating player enjoyment in games. *Computers in Entertainment (CIE)*, 3(3):3–3.

Tejeiro Salguero, R. and Morán, R. (2002). Measuring problem video game playing in adolescents. *Addiction*, 97(12):1601–1606.

ten Berge, M. and De Raad, B. (2002). The structure of situations from a personality perspective. *European Journal of Personality*, 16(2):81–102.

Thue, D., Bulitko, V., Spetch, M., and Wasylishen, E. (2007). Learning player preferences to inform delayed authoring. In *AAAI Fall Symposium on Intelligent Narrative Technologies*. AAAI Press, Arlington.

Thurstone, L. (1934). The vectors of mind. *Psychological Review*, 41(1):1.

Tychsen, A. (2008). Crafting user experience via game metrics analysis. In *Proceedings of the Workshop "Research Goals and Strategies for Studying User Experience and Emotion" at the 5th Nordic Conference on Human-computer interaction: building bridges (NordiCHI), Lund, Sweden*, pages 20–22.

van Aart, J., Bartneck, C., Hu, J., Rauterberg, M., and Salem, B. (2010). How to behave as alice in wonderland–about boredom and curiosity. *Entertainment Computing*, 1(3):125–137.

van Aart, J., Salem, B., Bartneck, C., Hu, J., and Rauterberg, M. (2008). Designing for experience: Arousing boredom to evoke predefined user behaviour. In *6th Design and emotion conference*, Hong Kong.

van den Herik, H. J. (1983). *Computerschaak, schaakwereld en kunstmatige intelligentie*. PhD thesis, Technische Hogeschool, Onderafdeling der Wiskunde en Informatica. Academic Service, Schoonhoven, The Netherlands.

van den Herik, H. J., Donkers, H. H. L. M., and Spronck, P. H. M. (2005). Opponent Modelling and Commercial Games. In *Proceedings of IEEE 2005 Symposium on Computational Intelligence and Games CIG05*, pages 15–25. IEEE.

van Lankveld, G., Schreurs, S., and Spronck, P. H. M. (2009). Psychologically verified player modelling. In Breitlauch, L., editor, *The 10th International Conference on Intelligent Games and Simulation (GAME-ON 2009)*, pages 12–19.

van Lankveld, G., Schreurs, S., Spronck, P. H. M., and van den Herik, H. J. (2011a). Extraversion in games. In van den Herik, H. J., Plaat, A., and Iida, H., editors, *CG'10 Proceedings of the 7th international conference on Computers and Games*, pages 263–275.

van Lankveld, G., Spronck, P. H. M., and Rauterberg, M. (2008). Difficulty scaling through incongruity. In *Proceedings of the Fourth Artificial Intelligence and Interactive Digital Entertainment Conference*, Stanford, California, USA.

van Lankveld, G., Spronck, P. H. M., van den Herik, H. J., and Arntz, A. R. (2011b). Games as personality profiling tools. In *2011 IEEE Conference on Computational Intelligence and Games (CIG)*, pages 197–202. IEEE.

van Lankveld, G., Spronck, P. H. M., van den Herik, H. J., and Rauterberg, M. (2010). Incongruity-based adaptive game balancing. In *12th International Conference, ACG 2009*, pages 208–221. Springer-Verlag, Heidelberg, Germany.

Wiggins, J. (1996). *The five-factor model of personality: Theoretical perspectives.* The Guilford Press.

Wong, W., Shen, C., Nocera, L., Carriazo, E., Tang, F., Bugga, S., Narayanan, H., Wang, H., and Ritterfeld, U. (2007). Serious video game effectiveness. In *Proceedings of the international conference on Advances in computer entertainment technology*, pages 49–55. ACM.

Yannakakis, G. and Hallam, J. (2007a). Game and player feature selection for entertainment capture. In *Proceedings of the IEEE Symposium on Computational Intelligence and Games*, pages 244–251.

Yannakakis, G. and Hallam, J. (2007b). Towards Optimizing Entertainment in Computer Games. *Applied Artificial Intelligence*, 21(10):933–971.

Yannakakis, G. N. and Hallam, J. (2008). Real-time adaptation of augmented-reality games for optimizing player satisfaction. *2008 IEEE Symposium On Computational Intelligence and Games*, pages 103–110.

Appendix A

Incongruity questionnaire

For the sake of clarity, the questions were translated from the original Dutch version into English (see Table A.1). The questions loaded on six categories: boredom, frustration, pleasure, concentration, curiosity, and not used. The symbol (i) indicates that the question was inverted to the category. Inverting the score leads to an appropriate category score.

Table A.1: The incongruity questionnaire

	Question	Category
1	In my opinion the game was user friendly	not used
2	I am interested in how the game works	curiosity
3	I had fun while playing the game	pleasure
4	I want to know more about the game	curiosity
5	I can easily concentrate on what I need to do during the game	concentration
6	The time passed quickly	boredom (i)
7	I got distracted	concentration (i)
8	I felt involved in the task	not used
9	The game frustrated me	frustration
10	The game made me curious	curiosity
11	I felt challenged	pleasure
12	The time passed slowly	boredom
13	The task fascinated me	curiosity
14	I was thinking about other things during play	concentration (i)
15	I found the game to be fun	pleasure
16	I was bored	boredom (i)
17	The game was tedious	boredom
18	I would like to ask questions about the game	curiosity
19	I thought the game was hard	frustration
20	I was alert during the game	concentration
21	I was day dreaming during the game	concentration (i)
22	I want to play the game again	pleasure
23	I understood what I was supposed to do in the game	not used
24	The game was easy	not used
25	I feel I was not doing well during the game	frustration
26	I was annoyed during play	frustration

Appendix B

Balancing glove

GLOVE's game world is quite bare, and there is little diversity in the challenges that the player faces. This is done on purpose. The aim of GLOVE is to provide the player with an entertaining experience, by only varying the number and types of enemies that the knight has to fight.

The game has three difficulty levels, named easy, balanced, and hard. While it is trivial to add more difficulty levels, for the present experiment these three were deemed sufficient. When the difficulty is set at easy, the game aims at having the knight win the game with about 50% of his health remaining. When the difficulty is set at hard, the game aims at having the knight lose the game when he has progressed through about 50% of the game world. When the difficulty is balanced, the game aims at having the knight experiencing a narrow victory or a narrow defeat. The game accomplishes this by controlling the number and types of spawned enemies. In this way the easy and hard levels try to keep incongruity stable and high, while the balanced level tries to keep incongruity stable and at a minimum.

For each enemy type, the game retains the average damage that the enemy type involved does to the knight. The number in which the average damage is expressed can be positive or negative (or zero). If positive, it means that the knight on average loses health due to an encounter with this enemy type. If negative, it means that the knight on average gains health due to an encounter with this enemy type. Gaining health is possible because the enemies leave health tokens upon dying, and it is certainly possible to kill an enemy without it being able to damage the knight.

Enemies are spawned just outside the knight's vision every 10 cells that the knight has progressed towards the right end of the world. The number and types of spawned enemies are determined by the game based on (1) the difficulty level, (2) the knight's progress, (3) the knight's health, and (4) the average damage done by each type of enemy.

The net result of the spawning procedure is that between 2 and N enemies are spawned, N being a number that depends on (1) the difficulty setting, and (2) the current progress of the knight. The spawned group of enemies is made up of the three types described above (dragon, ninja, and witch). The combined difficulty of the enemies is set so that, according to the players's past experience with these types of enemies, the knight is expected to lose or gain a predetermined amount of health determined by the difficulty.

Enemy spawn algoritm

Here, the algorithm used to spawn enemies is described. In brief, the enemy spawn algorithm checks how much health the player still has and how far along the level he has progressed. A calculation is made how much health the player still needs to lose based on the progress and current health. The amount of health to lose is then modified according to the difficulty setting. On easy mode the amount of health to lose is decreased while on hard mode the amount of health to lose is increased.

Once the amount of health to lose has been calculated the algorithm starts to spawn enemies. It continues to spawn enemies randomly until the total average damage for the spawned group of enemies is equal to the amount of health to lose. Once this goal has been reached the enemy spawning stops.

```
procedure spawnEnemies()
begin
 needed_health := getNeededHealth(getCurrentProgress()) +
    getModifier(getDifficulty());
 health_to_lose := getCurrentHealth() − needed_health;
 expected_health_loss := 0;
 spawned:=0;
 while (spawned < 2) or ((expected_health_loss < health_to_lose)
    and (spawned < getMaxSpawn( getLastSpawned()+1, getDifficulty
    ()))) do
 begin
  enemyType := spawnRandomEnemy(health_to_lose);
  health_to_lose := health_to_lose − getAverageDamage(enemyType);
  spawned := spawned + 1;
 end;
end;
```

An explaination of the various parts of the spawing algorithm is given below.

1. **getCurrentProgress()** returns a percentage that expresses how far the knight has progressed through the game world.

2. **getNeededHealth()**, which gets the knight's current progress as a parameter, returns the number of hitpoints that the knight needs to traverse the remaining part of the game world, if unobstructed by enemies.

3. **getDifficulty()** returns the difficulty level (easy, balanced, or hard).

4. **getModifier()**, which gets the difficulty level as a parameter, returns a number that is 50 for the easy difficulty level, 5 for balanced, and -50 for hard.

5. **getCurrentHealth()** returns the current health of the knight.

6. **health_to_lose** contains the number of hitpoints that the knight should lose for the game to reach the goal determined by the difficulty level. This number can be negative, which indicates that the knight should actually gain health.

7. **spawnRandomEnemy()** spawns an enemy. This function gets **health_to_lose()** as a parameter, by which it determines that either (1) it should spawn enemies that are likely to gain the knight some health, or (2) it should spawn enemies that cause the knight to lose health, or (3) it should spawn enemies that have a neutral influence on health. To avoid the algorithm getting into an endless loop, when the function should make the knight gain health, it will always allow dragons to be spawned; moreover, when it should make the knight lose health, it will always allow ninjas and witches to be spawned. It should be noted that with more enemies, it is harder to avoid damage; even if the player has reached a skill level in which he manages to gain health from all enemy types, he will consider the game harder if he gets surrounded by more of them.

8. **getLastSpawned()** returns the number of enemies that were spawned the last time.

9. **getMaxSpawn()** returns the maximum number that can be spawned. It gets two parameters, the first being a maximum that cannot be exceeded, and the second being the difficulty level, which it uses to determine a maximum number of enemies to spawn, which is 5 for easy, 7 for balanced, and 9 for hard.

10. **getAverageDamage()** is the last function. It returns the average damage done by the enemy type that is used as the parameter.

Difficulty algorithm

The algorithm to set the game difficulty is based on the player skill in dealing with the different types of enemies. The algorithm is given in Appendix C.

Appendix C

Glove difficulty algorithm

Described here is the Glove difficulty algorithm. The algorithm decides (1) which enemy type to spawn, and (2) how many of each type. In brief, it selects a random enemy from one of the three available types. For each enemy type, the average damage done to the player has been recorded and the algorithm selects an enemy appropriate to increasing or decreasing the average damage done to the player. The average damage done by each enemy is calculated by starting at -5 (because each defeated enemy drops a health token). For each time the enemy hits the player, the damage done for that enemy is added to the total. When the enemy is defeated the total is averaged with the totals of the last five enemies of that type to form the average. This mechanism means that defeating sufficient enemies that do no damage may result in a negative damage average. Enemies with a negative damage average may be used by the algorithm as a means for healing the player when his health is too low in comparison to his progress and the difficulty level.

```
function spawnRandomEnemy(health_to_lose)
begin
 while TRUE
 begin enemyType = getRandomEnemyType();
  if (health must be lost) then (only dragons are spawned)
  else
  if (they are likely to cost health when health_to_lose >= 0)
     then
  begin
   if (enemyType != DRAGON) or (getAverageDamage(enemyType) >= 0)
      then
   begin spawnEnemy(enemyType);
    return enemyType;
   end;
  end
  else
  if (health must be gained, ninjas and witches are only spawned
     if they are likely to supply health) then
  begin
   if (enemyType = DRAGON) or (getAverageDamage(enemyType) < 0)
      then
   begin
    spawnEnemy(enemyType);
    return enemyType;
   end;
  end;
 end;
end;
```

.

Appendix D

Incongruity contrast analysis

Here the contrast analysis belonging to the incongruity experiment of Chapter 3 is presented. Table D.1 shows the contrasts between the five emotions measured in the questionnaire in contrast to the three difficulty levels in the experiment (easy, balanced, hard). In the column 'difficulty', 'Level 1' represents the easy difficulty, 'Level 2' represents the balanced difficulty, and 'Level 3' represents the hard difficulty.

Table D.1: Contrasts

Source	Measure	Difficulty	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power
Difficulty	Boredom	Level 1 vs. Level 2	.001	1	.001	.002	.963	.000	.002	.050
		Level 2 vs. Level 3	.555	1	.555	.884	.360	.049	.884	.144
	Frustration	Level 1 vs. Level 2	23.054	1	23.054	8.811	.009	.341	8.811	.799
		Level 2 vs. Level 3	41.908	1	41.908	21.633	.000	.560	21.633	.992
	Pleasure	Level 1 vs. Level 2	.185	1	.185	.101	.755	.006	.101	.060
		Level 2 vs. Level 3	12.089	1	12.089	5.794	.028	.254	5.794	.622
	Concentration	Level 1 vs. Level 2	.138	1	.138	.173	.683	.010	.173	.068
		Level 2 vs. Level 3	.340	1	.340	.553	.467	.031	.553	.108
	Curiosity	Level 1 vs. Level 2	.009	1	.009	.018	.894	.001	.018	.052
		Level 2 vs. Level 3	.057	1	.057	.070	.794	.004	.070	.057

Level 1 = easy, level 2 = balanced, level 3 = hard.

Appendix E

Extraversion in games instruction booklet

The instruction booklet and the two questionnaires, the extraversion part of the NEO-PI-R and the general information questionnaire, are included here in the native language of the participants for the experiment (the Dutch language). Presented here is the "A" version of the booklet. In the "B" version, Sections 1 and 2 are reversed. The contents of Section 3 are last in both versions.

Bedankt voor uw deelname aan dit experiment.

Het onderzoek bestaat uit drie delen:

Als eerste krijgt u een vragenlijst met 48 vragen die beantwoord moeten worden met een score van 1 tot 5, waarbij 1 helemaal oneens is en 5 helemaal eens.

Na de vragenlijst zult u een kort computerspel gaan spelen waarbij het de bedoeling is dat u zich zo in het karakter verplaatst dat u de keuzes zou maken die u in het echte leven ook zou maken. Wanneer er in het spel verschillende gesprekskeuzes mogelijk zijn, is het ook de bedoeling dat u de keuze kiest die het meest lijkt op wat u in het normale leven zou zeggen.

Het spel is in het Engels, omdat het onderzoeksverslag ook in het Engels zal zijn. Als u grote moeite heeft met Engels kunt u dit aan de experimentleider laten weten.

Tot slot zult u een korte vragenlijst krijgen met een aantal simpele vragen.

Het onderzoek duurt in totaal ongeveer 30 tot 45 minuten.

De data verkregen uit de vragenlijsten en het computerspel zullen gebruikt worden voor onderzoek. Alle gegevens worden anoniem verwerkt.

Alvast bedankt voor uw medewerking.

Deel 1 - Vragenlijst

Het eerste deel bestaat uit een vragenlijst met 48 uitspraken. Lees eerst de instructies goed door en ga vervolgens naar de volgende pagina om aan de vragenlijst te beginnen.

Instructies:

Deze vragenlijst bevat 48 uitspraken. Lees elke uitspraak zorgvuldig en zet een cirkel rond het cijfer dat uw mening het beste weergeeft. Let erop dat u geen regels overslaat.

De scores betekenen het volgende:

1 = Helemaal oneens, 2 = Oneens, 3 = Neutraal, 4 = Eens, 5 = Helemaal eens

Zet dus een cirkel om het nummer dat u mening het beste weergeeft. Omcirkel slechts één mogelijkheid bij elke uitspraak.

Als u een antwoord wilt veranderen maak dan uw eerste keuze ongeldig door hier een kruis door te zetten en omcirkel alsnog het goede antwoord.

Voorbeeld:

Als u het bijvoorbeeld helemaal oneens bent met de uitspraak:

Ik zou wel een miljoen willen winnen

Dan vult u dat als volgt in:

1. Ik zou wel een miljoen willen winnen. (1) – 2 – 3 – 4 – 5

Wanneer u eerst aangeeft neutraal te zijn in de stelling, maar vervolgens toch van mening veranderd en vindt dat u er helemaal mee oneens bent, vult u het als volgt in:

1. Ik zou wel een miljoen willen winnen. (1) – 2 – X – 4 – 5

Ga naar de volgende pagina om aan de vragenlijst te beginnen.

	Oneens-Neutraal-Eens
1. Ik vind de meeste mensen die ik ontmoet echt aardig	1 – 2 – 3 – 4 – 5
2. Ik ga mensenmenigtes uit de weg	1 – 2 – 3 – 4 – 5
3. Ik ben dominant, krachtig en zelfverzekerd	1 – 2 – 3 – 4 – 5
4. Mijn stijl in werk en spel is kalm en ongehaast	1 – 2 – 3 – 4 – 5
5. Ik smacht vaak naar opwinding	1 – 2 – 3 – 4 – 5
6. Ik heb nog nooit letterlijk een gat in de lucht gesprongen van blijdschap	1 – 2 – 3 – 4 – 5
7. Ik vind het niet erg leuk om zomaar een praatje met iemand te maken	1 – 2 – 3 – 4 – 5
8. Ik houd er van veel mensen om me heen te hebben	1 – 2 – 3 – 4 – 5
9. Soms lukt het mij niet, genoeg voor mezelf op te komen	1 – 2 – 3 – 4 – 5
10. Mijn manier van doen is energiek en krachtig	1 – 2 – 3 – 4 – 5
11. Ik zou van een vakantie in Las Vegas niet genieten	1 – 2 – 3 – 4 – 5
12. Ik heb wel eens een intense vreugde of extase ervaren	1 – 2 – 3 – 4 – 5
13. Ik sta bekend als warm en vriendelijk mens	1 – 2 – 3 – 4 – 5
14. Ik geef er meestal de voorkeur aan om dingen alleen te doen	1 – 2 – 3 – 4 – 5
15. Ik heb vaak de leiding gehad in groepen waar ik bij hoorde	1 – 2 – 3 – 4 – 5
16. Mijn werkwijze is meestal traag maar gestadig	1 – 2 – 3 – 4 – 5
17. Ik heb wel eens dingen alleen maar voor de kick of de sensatie gedaan	1 – 2 – 3 – 4 – 5
18. Ik ben geen vrolijke optimist	1 – 2 – 3 – 4 – 5
19. Veel mensen vinden mij enigszins koel en afstandelijk	1 – 2 – 3 – 4 – 5
20. Ik heb echt behoefte aan gezelschap als ik lange tijd alleen ben	1 – 2 – 3 – 4 – 5
21. Tijdens bijeenkomsten laat ik anderen vaak het woord doen	1 – 2 – 3 – 4 – 5
22. Ik voel me vaak alsof ik barst van energie	1 – 2 – 3 – 4 – 5
23. Ik vermijd meestal films die schokkend of eng zijn	1 – 2 – 3 – 4 – 5
24. Soms loop ik over van geluk	1 – 2 – 3 – 4 – 5
25. Ik vind het echt leuk om met mensen te praten	1 – 2 – 3 – 4 – 5
26. Ik geef de voorkeur aan werk dat ik alleen kan doen zonder gestoord te worden	1 – 2 – 3 – 4 – 5
27. Anderen kijken vaak naar mij als er een beslissing genomen moet worden	1 – 2 – 3 – 4 – 5
28. Ik ben niet zo snel en levendig als anderen	1 – 2 – 3 – 4 – 5
29. Ik ben graag daar waar wat te beleven valt	1 – 2 – 3 – 4 – 5
30. Ik zie mezelf niet echt als een vrolijk en opgewekt persoon	1 – 2 – 3 – 4 – 5
31. Ik vind het gemakkelijk om vlot en plezierig met vreemden om te gaan	1 – 2 – 3 – 4 – 5
32. Ik zit met vakantie liever op een druk strand dan in een afgelegen hut in het bos	1 – 2 – 3 – 4 – 5
33. Ik ga liever mijn eigen gang dan dat ik leiding geef aan anderen	1 – 2 – 3 – 4 – 5
34. Het lijkt alsof ik altijd haast heb	1 – 2 – 3 – 4 – 5
35. Ik hou van de opwinding van de achtbaan	1 – 2 – 3 – 4 – 5
36. Ik ben een vrolijk en levendig iemand	1 – 2 – 3 – 4 – 5
37. Ik heb sterke emotionele banden met mijn vrienden	1 – 2 – 3 – 4 – 5
38. Feesten en partijtjes vind ik doorgaans vervelend	1 – 2 – 3 – 4 – 5
39. In conversaties ben ik doorgaans het meest aan het woord	1 – 2 – 3 – 4 – 5
40. Ik heb een jachtig leven	1 – 2 – 3 – 4 – 5
41. Ik voel me aangetrokken tot felle kleuren en opvallende kleding	1 – 2 – 3 – 4 – 5
42. Om mijn ervaringen te beschrijven gebruik ik zelden woorden als fantastisch of geweldig	1 – 2 – 3 – 4 – 5
43. Ik ben persoonlijk genteresseerd in de mensen met wie ik werk	1 – 2 – 3 – 4 – 5
44. Ik hou van feesten met veel mensen	1 – 2 – 3 – 4 – 5
45. Ik vind het niet gemakkelijk om de leiding op me te nemen	1 – 2 – 3 – 4 – 5
46. Ik ben een heel actief persoon	1 – 2 – 3 – 4 – 5
47. Ik maak graag deel uit van de menigte bij sportevenementen	1 – 2 – 3 – 4 – 5
48. Ik lach gemakkelijk	1 – 2 – 3 – 4 – 5

Deel 2 Computerspel

U zult nu een kort computerspel gaan spelen waarbij het de bedoeling is dat u zich in de hoofdpersoon verplaatst. U moet dus reageren zoals u in het echt zelf ook zou doen. Zorg er als u een keuze moet maken dus voor dat deze keuze u eigen gedrag het best uitdrukt.
Het spel zal beginnen in een korte testruimte, zodat u de besturing van het personage en de camera onder de knie kan krijgen. Vervolgens begint het experiment, wat ongeveer 15 tot 30 minuten duurt.

Instructies:

Aangezien dit onderzoek maar gebruik maakt van een klein deel van de mogelijkheden van het spel, hoeven niet alle functies uitgelegd te worden. De benodigde functies staan hieronder beschreven en worden in de testruimte ook nog eens uitgelegd. U kunt deze pagina voor u houden, zodat wanneer u iets vergeten bent u hier terug kan kijken. Als iets onduidelijk is aarzel niet om de experimentleider om opheldering te vragen.

De besturing:

Rondlopen:

Om rond te lopen moet u met de linker muisknop ergens op het scherm klikken. Het personage loopt dan naar de plek waar geklikt is.

Interactie met voorwerpen (computerspelers, deuren, kisten, etc.):

Om een interactie te starten moet u met de linker muisknop op een voorwerp of persoon klikken. Wanneer u bijvoorbeeld op een deur of een kist klikt zal u deze proberen open te maken. Wanneer u op een computerspeler klikt zal u proberen hier een conversatie mee te starten, of indien hij vijandig is zal u hem aanvallen.

Keuzes maken in een gesprek:

Wanneer u een gesprek heeft met een computerspeler zal linksboven in het scherm een kader worden geopend. Bovenin dat kader ziet u wat de computerspeler tegen u zegt. Onderin staan de reacties die u daarop kan geven. In de meeste gevallen kunt u maar 1 reactie kiezen, maar er zijn gesprekken waarin u meerdere reacties kunt kiezen. U kiest een reactie door hier met de linkermuisknop op te klikken.
De camera-hoek veranderen:

Het kan voorkomen dat u een deel van een ruimte niet goed kunt zien doordat de camera niet goed staat. U kunt de camera dan draaien door de muiscursor naar de zijkant van het scherm te brengen. U kunt ook de pijltjestoetsen op het toetsenbord gebruiken om de camera te draaien en de camera in en uit te zoomen. Met het muiswiel kunt u ook inzoomen en uitzoomen.

Voorwerpen oppakken en aantrekken:

U kunt een voorwerp oppakken door hier met de linker muisknop op te drukken. Wanneer u een voorwerp heeft opgepakt, komt deze in uw inventory terecht. Om deze inventory te openen klikt u op de i toets. Als u een voorwerp heeft wat aangetrokken kan worden, doet u dit door deze eerst aan te klikken in het inventory, en vervolgens te klikken op de plek in de inventory waar dat voorwerp gedragen kan worden.

Annuleren:

Tot slot kunt u altijd op escape drukken als er een ongewenst menu is geopend, dit menu wordt dan gesloten en de normale interface wordt weer hersteld.

Deel 3 Vragen achteraf

Mijn leeftijd is: .... jaar

Ik ben een: man / vrouw

De hoogste opleiding die ik heb afgerond is:
0 Lager onderwijs
0 Voorbereidend beroepsonderwijs (VBO, LTS, LHNO)
0 Algemeen vormend onderwijs (MAVO, HAVO, MULO)
0 Voorbereidend wetenschappelijk onderwijs (VWO, Gymnasium, Atheneum, HBS)
0 Middelbaar beroepsonderwijs (MBO, MTS, MEAO e.d.)
0 Hoger beroepsonderwijs (HEAO, HBO e.d.)
0 Universiteit (WO)

De ervaring die ik heb met computers zou ik beschrijven als:
0 Zeer weinig
0 Weinig
0 Normaal of gemiddeld
0 Veel
0 Zeer veel

De ervaring die ik heb met het spelen van computergames zou ik beschrijven als:
0 Zeer weinig
0 Weinig
0 Normaal of gemiddeld
0 Veel
0 Zeer veel

Mijn vaardigheid met de Engelse taal zou ik beschrijven als:
0 Zeer slecht
0 Slecht
0 Normaal of gemiddeld
0 Goed
0 Zeer goed

Ik vond de besturing van het spel:
0 Zeer lastig
0 Lastig
0 Normaal
0 Makkelijk
0 Zeer makkelijk

Ik vond de opdrachten en de voortgang in het spel:
0 Zeer onduidelijk
0 Onduidelijk
0 Normaal
0 Duidelijk
0 Zeer duidelijk

Kunt u beschrijven wat u deed in het computerspel toen u een minuut moest wachten?
....
....
....
Hartelijk bedankt voor uw medewerking!

Appendix F

NWN study descriptives

Table F.1: NWN study game variables: descriptive statistics

Variable Name	N	Range	Min	Max	Mean	Std. Dev
			Group 1			
TotalTime	44	3533	795	4328	2197.59	948.835
Moricho convend	44	1	0	1	.73	.451
convstart	44	60	7	67	31.68	16.133
convchoice	44	88	3	91	38.59	18.673
convend	44	49	7	56	27.98	11.927
Passed	44	248	32	280	109.70	55.781
combat	44	2	3	5	3.36	.574
death	44	0	3	3	3.00	.000
Kist	44	13	0	13	2.50	2.308
Sleutel	44	4	0	4	.84	.776
			Group 2			
Move Dream	44	19	13	32	17.91	4.826
Move First_Floor	44	16	0	16	3.91	3.402
Move My_House	44	40	0	40	7.41	8.740
Move Lumbercamp	44	64	0	64	14.16	14.832
Move Village	44	114	0	114	38.36	24.633
Move Rogerlson_house	44	12	0	12	3.25	2.721
Move Shop	44	22	0	22	2.41	3.872
Move Farm	44	17	0	17	3.52	4.311
Move Oldman_House	44	3	0	3	.41	.844
Move Inn	44	8	0	8	1.02	1.502
Move Inn_upstairs	44	9	0	9	1.07	1.993
Move Clearing	44	11	0	11	.55	1.970
Move Dark_forest	44	25	0	25	6.89	5.341
Move Strange_tower	44	15	0	15	2.64	3.491

Table F.1: NWN study game variables: descriptive statistics

Variable Name	N	Range	Min	Max	Mean	Std. Dev
Move Inside_the_tower	44	17	0	17	2.66	3.894
Move In_the_dark_cave	44	8	0	8	3.55	1.886
Group 3						
Con_Dream	44	8	7	15	10.25	1.296
Con_Lumbercamp	44	39	0	39	10.23	8.375
Con_Village	44	42	0	42	13.16	9.147
Con_Shop	44	33	0	33	6.80	6.144
Con_Clearing	44	21	0	21	2.39	4.652
Con_Rogerlson_house	44	21	0	21	5.57	5.479
Con_Inn	44	68	0	68	17.39	14.562
Con_Inside_the_tower	44	6	0	6	2.80	1.488
Con_In_the_dark_cave	44	4	0	4	2.20	1.357
Group 4						
Con MrRed	44	4	4	8	5.91	.741
Con MrBlue	44	5	2	7	4.34	.834
Con Siline	44	39	0	39	10.23	8.375
Con Marto	44	40	0	40	9.93	9.483
Con OudeMan	44	9	0	9	3.23	2.208
Con Dara	44	33	0	33	6.80	6.144
Con Maline	44	21	0	21	2.39	4.652
Con Boy	44	6	0	6	1.34	1.554
Con Evana	44	18	0	18	4.23	4.440
Con Serveerster	44	17	0	17	1.68	3.025
Con Barman	44	9	0	9	1.36	1.713
Con Jula	44	7	0	7	1.68	1.986
Con Krick	44	10	0	10	1.93	2.095
Con Bran	44	16	0	16	3.02	3.344
Con Burrick	44	12	0	12	3.41	2.936
Con Herbergier	44	20	0	20	4.30	3.903
Con Myztor	44	6	0	6	2.80	1.488
Con Morichio	44	4	0	4	2.20	1.357
Group 5						
Droom_Deur_Kerker	44	3	1	4	2.14	.795
Droom_Deur_Lange_Gang	44	3	0	3	1.07	.398
Droom_Deur_MrBlue_West	44	3	1	4	2.18	.657
Droom_Deur_MrBlue_Zuid	44	6	1	7	1.48	1.248
Droom_Deur_MrRed_Noord	44	6	1	7	1.43	1.228
Droom_Deur_MrRed_Oost	44	1	1	2	1.05	.211
Droom_Kerker_Halverwege	44	4	2	6	2.50	1.110
Droom_Kerker_Kist	44	4	0	4	.95	.963
Droom_Langegang_Halverwege	44	3	0	3	1.02	.340
Droom_Stropop	44	8	0	8	2.27	1.453

Table F.1: NWN study game variables: descriptive statistics

Variable Name	N	Range	Min	Max	Mean	Std. Dev
Droom_Uitgang	44	4	0	4	.20	.668
First_Bedroom	44	9	0	9	1.82	1.646
First_Guestroom	44	6	0	6	1.59	1.317
First_Stairs	44	7	0	7	.50	1.285
Home_Door_Front	44	11	0	11	1.18	2.170
Home_Door_Livingroom	44	10	0	10	2.20	2.258
Home_Livingroom_Full	44	5	0	5	1.11	1.450
Home_Livingroom_Halfway	44	12	0	12	2.27	2.434
Home_Stairs_Up	44	8	0	8	.64	1.753
Lumber_Behind_House	44	5	0	5	.75	1.260
Lumber_Front_Door	44	2	0	2	.07	.334
Lumber_Gate	44	12	0	12	2.34	3.011
Lumber_Halfway	44	15	0	15	3.36	3.505
Lumber_Mushrooms	44	4	0	4	.25	.781
Lumber_Path	44	29	0	29	6.43	6.403
Lumber_Tree	44	5	0	5	.95	1.346
Dorp_Achter_Huis_Oud	44	3	0	3	.57	.950
Dorp_Akker	44	6	0	6	1.05	1.329
Dorp_Boerderij_Achterom	44	5	0	5	1.25	1.314
Dorp_Boerderij_Bovenom	44	12	0	12	3.07	2.564
Dorp_Boerderij_Ingang	44	1	0	1	.18	.390
Dorp_Brug	44	21	0	21	5.23	4.097
Dorp_Herberg_Achteringang	44	2	0	2	.11	.387
Dorp_Herberg_Voordeur	44	2	0	2	.30	.594
Dorp_Huis_Oud	44	2	0	2	.07	.334
Dorp_Huis_Rog	44	2	0	2	.18	.495
Dorp_Meer_Noord	44	6	0	6	1.89	1.298
Dorp_Meer_Zuid	44	4	0	4	2.02	.876
Dorp_Pad_Naar_Bos	44	4	0	4	1.09	1.235
Dorp_Pad_Naar_Lumber	44	7	0	7	.82	1.618
Dorp_Pad_Openplek	44	2	0	2	.09	.421
Dorp_Rivier_Oost	44	8	0	8	1.39	2.305
Dorp_Rivier_West	44	5	0	5	1.16	1.430
Dorp_Stal	44	5	0	5	1.09	1.217
Dorp_Tuin_Oud	44	3	0	3	.55	.820
Dorp_Weg_Noord	44	30	0	30	8.41	6.739
Dorp_Weg_Zuid	44	25	0	25	7.43	6.075
Dorp_Winkel_Deur	44	2	0	2	.23	.522
Dorp_Zuid_Inham	44	2	0	2	.20	.594
Rogerlsons_Door_Front	44	6	0	6	.50	1.131
Rogerlsons_Door_Upstairs	44	1	0	1	.07	.255
Rogerlsons_Halfway	44	8	0	8	2.68	2.239

Table F.1: NWN study game variables: descriptive statistics

Variable Name	N	Range	Min	Max	Mean	Std. Dev
Winkel_Deur	44	4	0	4	.61	1.039
Winkel_Ver	44	21	0	21	1.80	3.488
Boerderij_Deur	44	3	0	3	.30	.668
Boerderij_Halverwege	44	10	0	10	2.14	2.611
Boerderij_Vol	44	6	0	6	1.09	1.537
Old_Man_Door_Front	44	0	0	0	.00	.000
Old_Man_Door_Upstairs	44	1	0	1	.05	.211
Old_Man_Halfway	44	2	0	2	.36	.750
Herberg_Achteringang	44	2	0	2	.32	.601
Herberg_Ingang	44	3	0	3	.45	.663
Herberg_Trap_Naar_Boven	44	5	0	5	.25	.918
Herberg_Boven_Kamer1	44	0	0	0	.00	.000
Herberg_Boven_Kamer2	44	2	0	2	.27	.694
Herberg_Boven_Kamer3	44	4	0	4	.66	1.160
Herberg_Boven_Trap	44	3	0	3	.14	.510
Clearing_Path_North	44	2	0	2	.14	.462
Clearing_Path_South	44	9	0	9	.41	1.575
Forest_Cave_Passage	44	2	0	2	.14	.510
Forest_North_Gate	44	11	0	11	2.55	1.910
Forest_North_Woods	44	3	0	3	1.05	.861
Forest_Road_Halfway	44	4	0	4	1.39	1.039
Forest_South_Gate	44	1	0	1	.02	.151
Forest_South_Woods	44	3	0	3	.34	.776
Forest_Temple_North	44	3	0	3	.48	.902
Forest_Temple_South	44	7	0	7	.32	1.095
Forest_Trees_Halfway	44	3	0	3	.61	.868
Tower_Entrance	44	2	0	2	.14	.409
Tower_Gate	44	8	0	8	.84	1.584
Tower_North	44	11	0	11	1.66	2.401
Toren_Bibliotheek	44	4	0	4	.61	1.017
Toren_Gang	44	6	0	6	.93	1.354
Toren_Slaapkamer	44	2	0	2	.27	.694
Toren_Studeerkamer	44	4	0	4	.57	1.043
Toren_Uitgang	44	5	0	5	.27	.872
DarkCave_Part1	44	0	0	0	.00	.000
DarkCave_Part2	44	5	0	5	.95	.746
DarkCave_Part3	44	2	0	2	.91	.473
DarkCave_Part4	44	3	0	3	.89	.618
DarkCave_Part5	44	2	0	2	.80	.462
MrRed_1 convstart	44	1	1	2	1.07	.255
MrRed_2 convstart	44	0	1	1	1.00	.000
MrRed_3 convstart	44	1	1	2	1.02	.151

Table F.1: NWN study game variables: descriptive statistics

Variable Name	N	Range	Min	Max	Mean	Std. Dev
MrRed_1.1 convchoice	44	2	0	2	.68	.601
MrRed_1.2 convchoice	44	1	0	1	.39	.493
MrRed_3.1 convchoice	44	2	0	2	.73	.499
MrBlue_1 convstart	44	2	0	2	1.02	.263
MrBlue_2 convstart	44	3	0	3	1.14	.554
MrBlue_3 convstart	44	2	0	2	1.18	.446
MrBlue_1.1 convchoice	44	1	0	1	.80	.408
MrBlue_1.2 convchoice	44	1	0	1	.20	.408
Siline_1 convstart	44	8	0	8	2.36	1.844
Siline_1.1 convchoice	44	4	0	4	1.02	.927
Siline_1.2 convchoice	44	5	0	5	1.14	1.112
Siline_1.3 convchoice	44	3	0	3	.25	.576
Siline_1.4 convchoice	44	5	0	5	1.20	1.133
Siline_1.1.2 convchoice	44	1	0	1	.41	.497
Siline_1.2.1 convchoice	44	2	0	2	.32	.518
Siline_1.2.2 convchoice	44	4	0	4	.82	.947
Siline_1.3.1 convchoice	44	1	0	1	.11	.321
Siline_1.3.2 convchoice	44	2	0	2	.14	.409
Siline_1.4.1 convchoice	44	1	0	1	.05	.211
Siline_1.4.2 convchoice	44	1	0	1	.09	.291
Siline_1.4.3 convchoice	44	1	0	1	.09	.291
Siline_1.3.2.1 convchoice	44	4	0	4	.95	.806
Siline_1.3.2.2 convchoice	44	2	0	2	.16	.479
Siline_1.4.2.1 convchoice	44	1	0	1	.23	.424
Siline_1.4.2.2 convchoice	44	3	0	3	.73	.694
Siline_1.4.3.1 convchoice	44	2	0	2	.09	.362
Siline_1.4.3.2 convchoice	44	2	0	2	.07	.334
Marto_1 convstart	44	14	0	14	3.59	3.598
Marto_1.1 convchoice	44	4	0	4	1.23	.859
Marto_1.2 convchoice	44	4	0	4	.59	1.064
Marto_1.3 convchoice	44	7	0	7	1.41	1.661
Marto_1.1.1 convchoice	44	2	0	2	.70	.668
Marto_1.1.2 convchoice	44	2	0	2	.43	.545
Marto_1.2.1 convchoice	44	1	0	1	.34	.479
Marto_1.2.2 convchoice	44	3	0	3	.23	.677
Marto_1.3.1 convchoice	44	3	0	3	.52	.902
Marto_1.3.2 convchoice	44	5	0	5	.89	1.185
OudeMan_1 convstart	44	4	0	4	1.57	.998
OudeMan_1.1 convchoice	44	3	0	3	1.00	.571
OudeMan_1.2 convchoice	44	1	0	1	.16	.370
OudeMan_1.3 convchoice	44	1	0	1	.02	.151
OudeMan_1.4 convchoice	44	1	0	1	.18	.390

Table F.1: NWN study game variables: descriptive statistics

Variable Name	N	Range	Min	Max	Mean	Std. Dev
OudeMan_1.5 convchoice	44	1	0	1	.07	.255
OudeMan_1.2.1 convchoice	44	0	0	0	.00	.000
OudeMan_1.2.2 convchoice	44	1	0	1	.16	.370
OudeMan_1.5.1 convchoice	44	1	0	1	.05	.211
OudeMan_1.5.2 convchoice	44	1	0	1	.02	.151
Dara_1 convstart	44	1	0	1	.98	.151
Dara_2 convstart	44	21	0	21	3.02	4.129
Dara_1.1 convchoice	44	1	0	1	.05	.211
Dara_1.2 convchoice	44	1	0	1	.75	.438
Dara_1.3 convchoice	44	1	0	1	.02	.151
Dara_2.1 convchoice	44	10	0	10	1.23	2.044
Dara_2.2 convchoice	44	2	0	2	.16	.479
Dara_2.3 convchoice	44	4	0	4	.59	.871
Maline_1 convstart	44	7	0	7	1.00	1.868
Maline_1.1 convchoice	44	7	0	7	.48	1.210
Maline_1.2 convchoice	44	5	0	5	.43	1.021
Maline_1.1.1 convchoice	44	7	0	7	.41	1.187
Maline_1.1.2 convchoice	44	1	0	1	.07	.255
Boy_1 convstart	44	6	0	6	1.34	1.554
Evana_1 convstart	44	6	0	6	1.66	1.725
Evana_1.1 convchoice	44	6	0	6	1.05	1.329
Evana_1.2 convchoice	44	2	0	2	.34	.568
Evana_1.3 convchoice	44	1	0	1	.09	.291
Evana_1.4 convchoice	44	1	0	1	.07	.255
Evana_1.1.1 convchoice	44	6	0	6	.66	1.140
Evana_1.1.2 convchoice	44	2	0	2	.34	.568
Evana_1.1.3 convchoice	44	1	0	1	.02	.151
Serveerster_1 convstart	44	17	0	17	1.68	3.025
Barman_1 convstart	44	6	0	6	.75	1.059
Barman_1.1 convchoice	44	1	0	1	.32	.471
Barman_1.2 convchoice	44	2	0	2	.30	.509
Jula_1 convstart	44	3	0	3	.66	.805
Jula_1.1 convchoice	44	2	0	2	.41	.583
Jula_1.2 convchoice	44	1	0	1	.20	.408
Jula_1.1.1 convchoice	44	2	0	2	.27	.499
Jula_1.1.2 convchoice	44	1	0	1	.09	.291
Jula_1.1.3 convchoice	44	1	0	1	.05	.211
Krick_1 convstart	44	4	0	4	.86	.930
Krick_1.1 convchoice	44	1	0	1	.20	.408
Krick_1.2 convchoice	44	1	0	1	.09	.291
Krick_1.3 convchoice	44	2	0	2	.23	.565
Krick_1.4 convchoice	44	1	0	1	.30	.462

Table F.1: NWN study game variables: descriptive statistics

Variable Name	N	Range	Min	Max	Mean	Std. Dev
Krick_1.1.1 convchoice	44	1	0	1	.02	.151
Krick_1.1.2 convchoice	44	1	0	1	.14	.347
Krick_1.1.3 convchoice	44	1	0	1	.05	.211
Krick_1.1.3.1 convchoice	44	1	0	1	.02	.151
Krick_1.1.3.2 convchoice	44	1	0	1	.02	.151
Bran_1 convstart	44	3	0	3	.91	.802
Bran_1.1 convchoice	44	3	0	3	.45	.697
Bran_1.2 convchoice	44	2	0	2	.41	.583
Bran_1.1.1 convchoice	44	2	0	2	.07	.334
Bran_1.1.2 convchoice	44	4	0	4	.59	.972
Bran_1.1.3 convchoice	44	0	0	0	.00	.000
Bran_1.1.4 convchoice	44	2	0	2	.23	.476
Bran_1.1.1.1 convchoice	44	0	0	0	.00	.000
Bran_1.1.1.2 convchoice	44	2	0	2	.05	.302
Bran_1.1.2.1 convchoice	44	1	0	1	.02	.151
Bran_1.1.2.2 convchoice	44	1	0	1	.14	.347
Bran_1.1.2.1.1 convchoice	44	1	0	1	.02	.151
Bran_1.1.2.1.2 convchoice	44	0	0	0	.00	.000
Bran_1.1.2.1.3 convchoice	44	0	0	0	.00	.000
Bran_1.1.2.2.1 convchoice	44	1	0	1	.09	.291
Bran_1.1.2.2.2 convchoice	44	1	0	1	.05	.211
Burrick_1 convstart	44	5	0	5	1.50	1.229
Burrick_1.1 convchoice	44	2	0	2	.57	.545
Burrick_1.2 convchoice	44	2	0	2	.48	.628
Burrick_1.3 convchoice	44	2	0	2	.34	.645
Burrick_1.4 convchoice	44	1	0	1	.07	.255
Burrick_1.2.1 convchoice	44	1	0	1	.07	.255
Burrick_1.2.2 convchoice	44	2	0	2	.39	.579
Herbergier_1 convstart	44	6	0	6	1.36	1.348
Herbergier_1.1 convchoice	44	2	0	2	.43	.625
Herbergier_1.2 convchoice	44	2	0	2	.27	.544
Herbergier_1.3 convchoice	44	1	0	1	.18	.390
Herbergier_1.4 convchoice	44	3	0	3	.61	.754
Herbergier_1.1.1 convchoice	44	1	0	1	.23	.424
Herbergier_1.1.2 convchoice	44	1	0	1	.18	.390
Herbergier_1.2.1 convchoice	44	0	0	0	.00	.000
Herbergier_1.2.2 convchoice	44	2	0	2	.25	.534
Herbergier_1.3.1 convchoice	44	1	0	1	.14	.347
Herbergier_1.3.2 convchoice	44	1	0	1	.05	.211
Herbergier_1.4.1 convchoice	44	1	0	1	.14	.347
Herbergier_1.4.2 convchoice	44	3	0	3	.45	.697
Myztor_1 convstart	44	1	0	1	.82	.390

Table F.1: NWN study game variables: descriptive statistics

Variable Name	N	Range	Min	Max	Mean	Std. Dev
Myztor_2 convstart	44	3	0	3	.36	.810
Myztor_1.1 convchoice	44	1	0	1	.14	.347
Myztor_1.2 convchoice	44	1	0	1	.61	.493
Myztor_1.3 convchoice	44	1	0	1	.02	.151
Myztor_1.1.1 convchoice	44	1	0	1	.77	.424
Myztor_1.1.2 convchoice	44	1	0	1	.07	.255
Myztor_1.1.3 convchoice	44	0	0	0	.00	.000
Moricho_1 convstart	44	2	0	2	.82	.540
Moricho_1.1 convchoice	44	1	0	1	.66	.479
Moricho_1.2 convchoice	44	1	0	1	.07	.255
Moricho_1.1.1 convchoice	44	1	0	1	.52	.505
Moricho_1.1.2 convchoice	44	1	0	1	.14	.347
Moricho_1.2.1.1 convchoice	44	0	0	0	.00	.000

Appendix G

NWN study coefficients

Table G.1: NWN study: regression coefficients openness.

	B	SE B	Beta
(Constant)	2.880	.323	
NW_Marto_1.3.2_convchoice	-.502	.046	-.349
NW_Barman_1.2_convchoice	-2.623	.146	-.784
NW_Krick_1.1.2_convchoice	2.386	.151	.486
NW_MrRed_1.1_convchoice	-1.155	.077	-.407
Con_Dream	.480	.034	.365
First_Guestroom	-.877	.049	-.678
First_Bedroom	.777	.049	.750
NW_Evana_1.1.1_convchoice	-.895	.059	-.598
NW_Maline_1.1_convchoice	.383	.054	.272
Forest_Temple_South	-.692	.048	-.445
Forest_South_Woods	.627	.062	.285
Lumber_Path	-.100	.009	-.377
Con_NW_Barman	.110	.041	.111
NW_Moricho_1.1.2_convchoice	.408	.143	.083
Tower_Entrance	-.768	.103	-.184
NW_Herbergier_1.1.2_convchoice	-.713	.133	-.163
Dorp_Herberg_Voordeur	.512	.080	.178
NW_Jula_1.1_convchoice	.540	.107	.185
Rogerlsons_Halfway	.203	.047	.267
NW_Dara_1.2_convchoice	.320	.120	.082
Droom_Stropop	-.079	.035	-.068

Table G.2: NWN study: regression coefficients conscientiousness.

	B	SE B	Beta
(Constant)	.125	.477	
NW_Evana_1_convstart	.776	.049	.661
Dorp_Winkel_Deur	-1.713	.136	-.442
NW_Bran_1.1.1_convchoice	1.989	.198	.328
NW_Barman_1.1_convchoice	-2.003	.177	-.466
Trigger_Shop	-.282	.020	-.540
NW_Krick_1.1.3.1_convchoice	4.582	.435	.341
Lumber_Gate	.332	.028	.494
Rogerlsons_Door_Front	-.804	.076	-.449
Droom_Deur_MrRed_Oost	4.380	.477	.456
NW_Marto_1.3.1_convchoice	.661	.086	.294
NW_Marto_1.2_convchoice	-.562	.086	-.295
NW_Siline_1.4.3_convchoice	1.427	.318	.205
Lumber_Behind_House	-.331	.078	-.206
Dorp_Meer_Noord	.208	.052	.133
Kist	-.131	.030	-.149
NW_Moricho_1.1.2_convchoice	-.633	.206	-.109
NW_MrRed_3.1_convchoice	-.584	.166	-.144
NW_OudeMan_1.2.2_convchoice	-.624	.236	-.114

Table G.3: NWN study: regression coefficients extraversion.

	B	SE B	Beta
(Constant)	10.822	1.513	
NW_Barman_1.1_convchoice	-2.307	.394	-.496
NW_Burrick_1.3_convchoice	3.237	.419	.953
Boerderij_Deur	-2.093	.362	-.638
Dorp_Winkel_Deur	1.778	.366	.424
Dorp_Boerderij_Achterom	-.606	.15	-.363
NW_Bran_1.1.2.2.1_convchoice	-2.671	.773	-.354
NW_Siline_1.2.1_convchoice	-.902	.356	-.213
Con_NW_MrRed	-.605	.254	-.205

Table G.4: NWN study: regression coefficients agreeableness.

	B	SE B	Beta
(Constant)	4.846	.093	
NW_Siline_1.1.2_convchoice	2.445	.021	.678
Toren_Bibliotheek	-.712	.013	-.404
Tower_Gate	.450	.006	.398
NW_Dara_1.1_convchoice	-5.650	.059	-.664
Droom_Deur_MrBlue_West	.782	.019	.286
NW_Krick_1.2_convchoice	-.926	.036	-.150
NW_MrBlue_3_convstart	-1.600	.021	-.398
NW_OudeMan_1.3_convchoice	-1.472	.046	-.124
Herberg_Boven_Kamer2	-.917	.012	-.355
Forest_South_Gate	-4.624	.069	-.389
combat	-.904	.013	-.290
NW_Burrick_1.2.2_convchoice	-1.186	.017	-.383
NW_Herbergier_1.3.1_convchoice	.389	.027	.075
NW_Herbergier_1.4.2_convchoice	.714	.018	.278
Forest_Cave_Passage	.466	.014	.133
Trigger_Dream	.107	.002	.287
Tower_Entrance	.889	.015	.203
NW_Evana_1.4_convchoice	-1.093	.041	-.155
NW_Myztor_2_convstart	-.346	.010	-.156
Winkel_Deur	.209	.009	.121
NW_Siline_1.2.2_convchoice	-.247	.011	-.130
Droom_Kerker_Kist	.149	.010	.080
NW_OudeMan_1_convstart	-.184	.011	-.103
Forest_North_Gate	.100	.005	.106
NW_Siline_1.3.1_convchoice	.222	.026	.040
NW_Herbergier_1.2.2_convchoice	-.354	.022	-.105
DarkCave_Part5	-.185	.026	-.048
NW_Bran_1.1.2_convchoice	.119	.012	.065
NW_Maline_1.1_convchoice	.082	.008	.056
Winkel_Ver	-.019	.003	-.036
Dorp_Meer_Noord	.056	.010	.040
Dorp_Winkel_Deur	-.054	.014	-.016

Table G.5: NWN study: regression coefficients neuroticism.

	B	SE B	Beta
(Constant)	8.430	.302	
Tower_Gate	.560	.029	.464
NW_Myztor_1_convstart	-1.158	.080	-.236
NW_Siline_1.4.2.1_convchoice	-2.254	.102	-.500
Dorp_Rivier_West	-.728	.026	-.544
TotalTime	.002	.000	.810
NW_Bran_1_convstart	-1.426	.053	-.598
NW_Jula_1.1.2_convchoice	-3.000	.199	-.456
NW_Burrick_1.2.2_convchoice	.742	.070	.225
Trigger_Inn	-.508	.033	-.399
NW_Dara_1.1_convchoice	-1.125	.138	-.124
Forest_South_Woods	-.953	.046	-.387
Droom_Deur_MrBlue_Zuid	.256	.028	.167
Lumber_Front_Door	1.054	.124	.184
NW_MrRed_3_convstart	-3.712	.254	-.293
Herberg_Ingang	1.003	.068	.348
NW_Krick_1.1.2_convchoice	-1.373	.104	-.249
Forest_Trees_Halfway	.284	.035	.129
NW_Herbergier_1.3.1_convchoice	-1.042	.107	-.189
NW_Barman_1.1_convchoice	.649	.082	.160
NW_Marto_1.3.1_convchoice	.234	.038	.111
NW_Jula_1.1.3_convchoice	-.960	.185	-.106
NW_Bran_1.1.2.2.2_convchoice	-.561	.148	-.062
Trigger_Inn_upstairs	-.058	.022	-.061
Con_NW_OudeMan	.043	.019	.049

Appendix H

NWN study significant correlations

Table H.1 contains the group numbers, variablenames, correlations, significance levels and participant numbers, and traits for the experiments of Chapter 5. Only significant correlations are shown. Correlations with * have a significance level of $p < 0.05$ and correllations with ** have a significance level of $p < 0.01$. The following abbreviations apply: O = Openness, C = Conscientiousness, E = Extraversion, A = Agreeableness, and N = Neuroticism.

Table H.1: NWN study: significant correlations

	O	C	E	A	N
		Group 1			
TotalTime	-.105	.273*	.051	.101	-.450**
Moricho convend	-.086	-.174	-.204	-.039	.281*
convstart	.055	-.034	.240	-.017	-.505**
convchoice	-.007	-.049	.241	-.007	-.490**
convend	.081	-.075	.225	-.046	-.508**
Passed	-.056	.168	.067	.064	-.439**
		Group 2			
Move Dream	.293*	-.027	.207	-.082	-.031
Move First_Floor	-.092	.204	-.213	.105	-.401**
Move Lumbercamp	.055	.045	.208	.077	-.401**
Move Village	-.073	.199	.092	.070	-.464**
Move Farm	-.026	.006	-.026	.091	-.258*
Move Inn	-.057	.103	-.055	-.005	-.264*
Move In_the_dark_cave	.002	-.207	-.201	-.050	.293*
		Group 3			
Con_Lumbercamp	-.053	.135	-.001	.047	-.448**
Con_Village	-.091	.223	.015	.127	-.530**
Con_Shop	-.133	.057	.158	.012	-.355**

Table H.1: NWN study: significant correlations

	O	C	E	A	N
Con_Clearing	-.098	-.204	.313*	-.117	.022
Con_Rogerlson_house	.035	-.196	.371**	-.021	-.280*
Con_Inn	.213	-.163	.266*	-.069	-.369**
Group 4					
Con Siline	-.053	.135	-.001	.047	-.448**
Con Marto	-.089	.208	.025	.182	-.513**
Con OudeMan	.004	.033	-.044	-.255*	.011
Con Dara	-.133	.057	.158	.012	-.355**
Con Maline	-.098	-.204	.313*	-.117	.022
Con Boy	.125	-.035	.098	-.038	-.424**
Con Evana	.000	-.230	.423**	-.013	-.197
Con Serveerster	.283*	-.088	.180	-.082	-.274*
Con Barman	-.135	.158	-.074	-.032	-.417**
Con Jula	.078	-.103	.281*	.058	-.108
Con Krick	.199	-.163	.229	-.081	-.371**
Con Bran	.244	-.274*	.308*	-.072	-.216
Con Burrick	.285*	-.179	.182	-.188	-.220
Con Herbergier	.064	-.099	.220	.036	-.378**
Group 5					
Move Dream	.293*	-.027	.207	-.082	-.031
Move First_Floor	-.092	.204	-.213	.105	-.401**
Move Lumbercamp	.055	.045	.208	.077	-.401**
Move Village	-.073	.199	.092	.070	-.464**
Move Farm	-.026	.006	-.026	.091	-.258*
Move Inn	-.057	.103	-.055	-.005	-.264*
Move In_the_dark_cave	.002	-.207	-.201	-.050	.293*
MrRed_1 convstart	-.014	.005	.056	-.148	-.271*
MrRed_3 convstart	-.078	-.105	-.095	.031	.269*
MrRed_1.1 convchoice	.046	.063	.067	-.131	-.378**
MrRed_1.2 convchoice	-.063	-.075	-.052	.084	.322*
MrBlue_1 convstart	-.327*	.125	-.011	.314*	-.053
MrBlue_2 convstart	.294*	-.171	.052	-.160	.046
Siline_1 convstart	-.062	.114	.006	.069	-.469**
Siline_1.1 convchoice	.067	-.043	.046	.201	-.295*
Siline_1.2 convchoice	.013	.145	.036	.049	-.456**
Siline_1.3 convchoice	-.041	.079	-.155	-.113	-.267*
Siline_1.4 convchoice	-.113	.176	-.003	-.020	-.376**
Siline_1.1.2 convchoice	-.086	-.057	.150	.458**	-.094
Siline_1.2.2 convchoice	-.001	.146	.061	.042	-.415**
Siline_1.3.1 convchoice	-.052	.020	-.224	-.129	-.260*
Siline_1.4.3 convchoice	-.199	.202	.040	.109	-.333*
Siline_1.3.2.1 convchoice	.069	.115	-.007	-.092	-.287*

Table H.1: NWN study: significant correlations

	O	C	E	A	N
Siline_1.3.2.2 convchoice	-.261*	.100	-.018	.042	-.204
Siline_1.4.2.1 convchoice	.247	-.171	.339*	-.134	-.265*
Siline_1.4.3.2 convchoice	-.265*	.150	-.060	.198	-.166
Marto_1 convstart	-.097	.214	.027	.189	-.484**
Marto_1.1 convchoice	-.113	.241	-.221	.115	-.417**
Marto_1.2 convchoice	-.129	.129	.038	.042	-.431**
Marto_1.3 convchoice	.000	.151	.142	.231	-.488**
Marto_1.1.1 convchoice	-.358**	.234	-.340*	.025	-.096
Marto_1.1.2 convchoice	.211	.120	.005	.117	-.386**
Marto_1.2.1 convchoice	-.148	.091	.126	.148	-.294*
Marto_1.2.2 convchoice	-.096	.127	-.025	-.026	-.407**
Marto_1.3.1 convchoice	.052	.124	.220	.350**	-.174
Marto_1.3.2 convchoice	-.040	.118	.031	.057	-.551**
Dara_2 convstart	-.111	.046	.172	.033	-.321*
Dara_1.2 convchoice	-.067	.118	-.269*	.148	-.023
Dara_2.1 convchoice	-.146	.113	.199	.106	-.342*
Maline_1 convstart	-.040	-.182	.277*	-.097	-.022
Maline_1.1 convchoice	-.231	-.193	.349*	-.122	.152
Maline_1.1.1 convchoice	-.286*	-.177	.305*	-.125	.214
Maline_1.1.2 convchoice	.236	-.090	.237	.005	-.271*
Boy_1 convstart	.125	-.035	.098	-.038	-.424**
Evana_1 convstart	.029	-.216	.458**	-.041	-.219
Evana_1.1 convchoice	-.002	-.216	.290*	.027	-.124
Evana_1.2 convchoice	-.051	-.138	.389**	-.127	-.032
Evana_1.3 convchoice	.056	-.008	.198	-.024	-.380**
Evana_1.4 convchoice	-.056	.005	.191	.005	-.325*
Evana_1.1.2 convchoice	.192	-.180	.268*	-.012	-.296*
Evana_1.1.3 convchoice	.133	-.024	.057	-.141	-.274*
Serveerster_1 convstart	.283*	-.088	.180	-.082	-.274*
Barman_1 convstart	-.088	.129	-.068	-.049	-.396**
Barman_1.1 convchoice	-.374**	.255*	-.280*	.030	-.038
Barman_1.2 convchoice	.073	.028	.152	-.032	-.543**
Jula_1.1 convchoice	.127	-.173	.345*	-.055	-.010
Jula_1.2 convchoice	-.079	.129	-.091	.263*	-.274*
Jula_1.1.1 convchoice	.077	-.111	.299*	.035	.047
Jula_1.1.2 convchoice	.166	-.175	.079	-.292*	-.004
Krick_1 convstart	.145	-.081	.142	-.058	-.467**
Krick_1.1 convchoice	.155	-.199	.303*	-.023	.127
Krick_1.2 convchoice	.166	-.091	.119	-.247	-.380**
Krick_1.4 convchoice	.104	-.022	-.006	.245	-.482**
Krick_1.1.3 convchoice	.291*	-.323*	.354**	-.263*	-.003
Krick_1.1.3.1 convchoice	.203	-.266*	.133	-.227	.179

Table H.1: NWN study: significant correlations

	O	C	E	A	N
Krick_1.1.3.2 convchoice	.203	-.185	.362**	-.141	-.183
Bran_1 convstart	.191	-.240	.315*	.106	-.407**
Bran_1.1 convchoice	.270*	-.260*	.329*	-.107	-.087
Bran_1.2 convchoice	-.110	-.007	-.069	.257*	-.337*
Bran_1.1.1 convchoice	.148	-.178	.421**	.004	-.125
Bran_1.1.2 convchoice	.252*	-.246	.230	-.207	-.050
Bran_1.1.1.2 convchoice	.203	-.185	.362**	-.141	-.183
Bran_1.1.2.2.2 convchoice	.089	-.265*	.082	-.017	.256*
Burrick_1 convstart	.263*	-.153	.201	-.158	-.239
Burrick_1.1 convchoice	-.191	.281*	-.237	-.022	.086
Burrick_1.2 convchoice	.281*	-.198	.123	-.193	-.162
Burrick_1.3 convchoice	.350**	-.329*	.254*	-.132	-.134
Burrick_1.4 convchoice	.152	-.090	.417**	.005	-.378**
Burrick_1.2.1 convchoice	.152	-.185	-.034	.259*	.050
Burrick_1.2.2 convchoice	.239	-.147	.134	-.354**	-.221
Herbergier_1 convstart	.104	-.097	.239	.008	-.459**
Herbergier_1.3 convchoice	.084	-.011	.088	-.302*	-.251
Herbergier_1.1.1 convchoice	.322*	-.314*	.122	.172	-.297*
Herbergier_1.3.1 convchoice	.163	-.062	.083	-.367**	-.123
Herbergier_1.3.2 convchoice	-.112	.081	.027	.045	-.262*
Myztor_1.1.1 convchoice	-.022	-.173	.014	-.081	.265*

Appendix I

Fallout instruction sheet

Interaction in computer games
Je doet mee aan een onderzoek over interactie in computerspelen. Het spel dat je zo meteen gaat spelen is Fallout 3. Je kunt alleen met het onderzoek meedoen als je dit spel niet eerder gespeeld hebt. Hoewel je gedurende dit onderzoek het spel slechts kort zult spelen, is het de bedoeling dat je het spel speelt met een houding alsof je het uiteindelijk helemaal uit zult spelen.

Lees deze tekst helemaal door voordat je gaat spelen.

Je zult in totaal +/- 45 minuten krijgen om vier fases van het spel te spelen. Gedurende de eerste twee fasen van het spel kun je de onderzoeker vragen stellen als iets onduidelijk is. Gedurende de derde en vierde fase mag je geen vragen stellen. De fasen worden hieronder uitgelegd.

De onderzoeker start het spel voor je op. Het spel begint met een filmpje. Als je er geen behoefte aan hebt dit filmpje te zien, kun je het afbreken door op *Escape* te drukken.

Fase 1: geboorte

In de eerste fase krijg je de geboorte van je spelkarakter te zien. Er wordt gevraagd een keuze te maken betreft het geslacht van het karakter en een naam te verzinnen. Hierna moet je het uiterlijk van het karakter bepalen. Dit uiterlijk is niet belangrijk voor het onderzoek, dus je kunt op “NEXT” klikken totdat je hiermee klaar bent (in het spel zou je eventueel later het uiterlijk nog kunnen veranderen).

Fase 2: 1 jaar oud

In de tweede fase wordt uitgelegd hoe je door de spelomgeving loopt, en hoe je met voorwerpen interacteert. Oefen hiermee totdat je er handigheid in hebt gekregen. Als je hier problemen mee hebt, vraag dan de onderzoeker het even uit te leggen. Om verder te kunnen komen moet je eerst een poortje openen om vervolgens het boekje dat op de grond ligt openen en naar de laatste pagina gaan. Daar moet je wat specialiteiten ophogen totdat alle punten verdeeld zijn. De precieze verdeling is niet belangrijk omdat je als speler later in het spel nog de mogelijkheid krijgt de specialiteiten te wijzigen. Als je dit hebt gedaan, wacht je tot "je vader" terugkeert. Luister naar hem en volg hem totdat de tweede fase eindigt.

Fase 3: 10 jaar oud

Je kunt vanaf nu geen vragen meer stellen aan de onderzoeker. In deze fase wordt de tiende verjaardag van je spelkarakter gevierd. Je kunt met diverse personen spreken. Let op: bij sommige personen heb je meer dan drie conversatiekeuzes; door op de pijltjes links in het conversatiescherm te klikken kun je meerdere opties zien.

Fase 4: 16 jaar oud

Als deze fase afgelopen is verschijnt er een boodschap op het scherm waarin gevraagd wordt te stoppen met spelen, ofwel het keyboard los te laten en de onderzoeker erbij te roepen. Stop dan onmiddellijk. Als je na 45 minuten nog niet klaar bent met de vier fases, dan is dit niet erg. De onderzoeker zal je dan vragen te stoppen wanneer de tijd om is.

Op de volgende pagina staat de belangrijkste besturing van het spel weergegeven.

De belangrijkste besturing van het spel is als volgt:

Gebruik de muis om om je heen te kijken.

W: Loop naar voren.
E: Interacteer met een voorwerp, b.v. open een deur of start een gesprek.
R: Starten of stoppen met vechten.

Je gebruikt de linkermuisknop om menukeuzes te maken, en, in een gevecht, te slaan of te schieten.

Hiermee kun je alles doen wat in het spel nodig is.

Daarnaast kun je eventueel ook nog de volgende besturingelementen gebruiken:

S: Loop achteruit
A: Loop naar links
D: Loop naar rechts
Spatie: Spring
Tab: Open de PIP-boy (vanaf 10 jaar)
V: Opent VATS mode (tijdens een gevecht)

Je gebruikt de rechtermuisknop om tijdens een gevecht met een schietwapen in te zoomen.

Appendix J

Demographics questionnaire

Hieronder volgen een aantal afsluitende vragen. Deze enqute is anoniem en wordt enkel gebruikt voor onderzoeksdoeleinden.

Mijn leeftijd is: ... jaar

Ik ben een: man / vrouw

De hoogste opleiding die ik heb afgerond is:
0 Lager onderwijs
0 Voorbereidend beroepsonderwijs (VBO, LTS, LHNO)
0 Algemeen vormend onderwijs (MAVO, HAVO, MULO)
0 Voorbereidend wetenschappelijk onderwijs (VWO, Gymnasium, Atheneum, HBS)
0 Middelbaar beroepsonderwijs (MBO, MTS, MEAO e.d.)
0 Hoger beroepsonderwijs (HEAO, HBO e.d.)
0 Universiteit (WO)

De ervaring die ik heb met computers zou ik beschrijven als:
0 Zeer weinig
0 Weinig
0 Normaal of gemiddeld
0 Veel
0 Zeer veel

De ervaring die ik heb met het spelen van computergames zou ik beschrijven als:
0 Zeer weinig
0 Weinig
0 Normaal of gemiddeld
0 Veel
0 Zeer veel

Mijn vaardigheid met de Engelse taal zou ik beschrijven als:
0 Zeer slecht
0 Slecht
0 Normaal of gemiddeld
0 Goed
0 Zeer goed

Ik vond de besturing van het spel:
0 Zeer lastig
0 Lastig
0 Normaal
0 Makkelijk
0 Zeer makkelijk

Ik vond de voortgang in het spel:
0 Zeer onduidelijk
0 Onduidelijk
0 Normaal

Appendix K

Fallout study dialog options

Table K.1 contains all the dialog text and dialog options found in the introduction of FALLOUT 3 and their respective variable names. Each line of the table shows the dialog text that is shown on screen followed by the response option available to the player. For each conversation text there are multiple response options available and the player must choose one in order to continue the conversation.

Table K.1: Fallout study: conversation dialog

variable	conversation text	conversation option
amac01		He's lucky Gomez stopped the fight before I really hurt him.
amac02		Nothing. Just wishing me a happy birthday.
amac03		The jerk tried to steal my sweetroll.
amac04		It was kind of my fault. You know how easy it is to make Butch mad...
amac05		He wanted my sweetroll, but I told him off. (Lie.)
amac06		He, um... He made me give him my sweetroll.
amac07		I don't want to talk about it.
amac08		Don't worry about me. I'm not scared of him.
amac09		I wouldn't give him what he wanted. He says he's going to make me sorry later.
amac10		Your dad's the Overseer. Can't you talk to him about Butch?

Table K.1: Fallout study: conversation dialog

variable	conversation text	conversation option
amac11		You didn't fool me. I just pretended not to know. (Lie.)
amac12		Great party, Amata! Thanks for doing this for me.
amac13		Is this it? Or hasn't the real party started yet?
amac14		I really have no idea...
amac15		Um... a date with Christine Kendall?
amac16		Just give me the present already!
beac01		Yeah, my Dad and Amata threw me a great party, didn't they?
beac02		Why are you talking to me like I'm five?
beac03		Yes. Can I have my present now?
beac04		Thank you. I will treasure it always. Is that all?
beac05		Um... thanks. I guess. Did you get me anything else?
beac06		A poem? You've got to be kidding me.
butc01		Mrs. Palmer said I didn't have to share. Since it's my birthday...
butc02		How about we share it? Half for me, half for you. That's fair, right?
butc03		You can have it. I don't even like sweetrolls.
butc04		Sure, Butch. (Spit on the sweetroll and give it to him.)
butc05		Go soak your head, Butch. I'm not giving you my sweetroll.
butc06		You do look hungry. What, your mom drank up all the ration coupons again?
butc07		I threw it away. But you're welcome to eat it off the floor if you want.
butc08		Mmmm... It sure was good when I ate it a few minutes ago.
dadc01		Don't worry about it, sir. It was nothing.

Table K.1: Fallout study: conversation dialog

variable	conversation text	conversation option
dadc02		He tried to take my birthday present!
dadc03		Really, I was the one who started it. It wasn't his fault.
dadc04		Here? We can't shoot a gun here.
dadc05		What do I get to kill?
dadc06		I'm out of BBs.
dadc07		Nothing's wrong, Dad.
dadc08		What's a Radroach?
dadc09		How do I kill it?
dadc10		I can do it!
dadc11		I can't do it!
dadc12		What kind of surprise?
dadc13		I don't like surprises.
dadc14		This is so great! Thanks!
dadc15		Thanks, Dad.
dadc16		Whatever. A shotgun would have been nice.
jonc01		I'm not a kid! I'm ten years old!
jonc02		Can it, Jonas. Where's my surprise present already?
jonc03		Oh... But Dad told me it was okay to come down here.
jonc04		Thanks a lot, Jonas. This is really cool.
jonc05		Couldn't you get me a real gun? A BB Gun is kind of lame.
jonc06		I guess so. I'm just not that into shooting things.
olpc01		Yes, ma'am.
olpc02		Oh, you didn't have to bring me a present, Mrs. Palmer.
olpc03		I hope you brought me something better than last year.
ovec01		Of course she likes me. I'm a really charming guy.
ovec02		She did a great job. But couldn't you have helped out a bit more?
stac01		It's really cool. Did you fix it up for me?
stac02		It's all right. Seems kind of old, though.

Table K.1: Fallout study: conversation dialog

variable	conversation text	conversation option
stac03		I hate it. How do I get it off?
stac04		Wally said my Pip-Boy was a piece of junk!
stac05		Thanks, Stanley.
stac06		Sure, whatever. Did you bring me anything for my birthday?
stac07		(Yawn) That's really really interesting.
secc01		Don't worry about it, sir. It was nothing.
secc02		He tried to take my birthday present!
secc03		Really, I was the one who started it. It wasn't his fault.
amac17		I'll see if I can talk some sense into them.
amac18		You have to fight your own battles. I can't help you with everything...
amac19		Look, Amata, this is none of my business. I don't want to get involved...
amac20		Tunnel Snakes? You guys are some kind of gang, is that it?
butc09		Leave her alone, or you'll answer to me.
butc10		Maybe I can help. She's very sensitive about her weight...
butc11		I'll just be going now.
butc12		That's it, Butch. You and me. Right now.
butc13		If you keep messing with her, the Overseer is going to come down on your gang.
butc14		CG03ButchQuitTopic
butc15		What's going on here?
butc16		Looks like you're having fun.
butc17		Hey, it's none of my business.
dadc17		If you say so, Dad.
dadc18		But, I'm sick. Really. (Lie.)
dadc19		Anything I need to know about the G.O.A.T.?

Table K.1: Fallout study: conversation dialog

variable	conversation text	conversation option
dadc20		Do we have to die in the Vault? Can't we ever leave?
dadc21		Is it true, Dad? Was everyone born in the Vault?
dadc22		Can we talk about, you know... mom?
broc01		Wow! That's what I've always wanted to be. My dreams are finally coming true!
broc02		Whatever. I just answered randomly. Is that how you got stuck with your job?
broc03		That can't be right! The stupid test got it all wrong!
broc04		Ohh... I'm feeling kind of sick, Mister B. Guess I'll have to reschedule...
broc05		Come on. I don't really have to take this stupid test, do I?
broc06		Sure, I'm ready. I bet I'll ace it!
broc07		Cool. Let me see the results and I'll fill it out myself.
broc08		The Overseer's bullshit makes my head spin. Can't you do it for me?
broc09		I think I'd rather just take the G.O.A.T. myself and take my chances.
broc10		I love using the computers and talking to my father's patients in the clinic.
broc11		Well, I shoot my BB Gun any chance I get. I can fix that thing blindfolded, too.
broc12		Look. I like blowing stuff up. I just love that... kaboom! Ya know?
broc13		Mr. B, if I told you what my interests are, youd have me locked up.
pauc01		Why do you listen to Butch?

Table K.1: Fallout study: conversation dialog

variable	conversation text	conversation option
pauc02		Hold on. You're in the Tunnel Snakes? You guys really do rule!
stac08		Tell me about the rebels.
stac09		What happened down here?
walc01		Tunnel Snakes? You guys are some kind of gang, is that it? / If you say so.
walc02		I've heard you do anything he tells you. And I mean anything.
walc03		Oh, I get it! Butch is the leader. You're just... a follower.
walc04		Hey, it's none of my business.
question1a	You are approached by a frenzied vault scientist, who yells, "I'm going to put my quantum harmonizer in your photonic resonation chamber!" What's your response?	But doctor, wouldn't that cause a parabolic destabilization of the fission singularity?
question1b	You are approached by a frenzied vault scientist, who yells, "I'm going to put my quantum harmonizer in your photonic resonation chamber!" What's your response?	Yeah? Up yours too, buddy!
question1c	You are approached by a frenzied vault scientist, who yells, "I'm going to put my quantum harmonizer in your photonic resonation chamber!" What's your response?	Say nothing, but grab a nearby pipe and hit the scientist in the head to knock him out. For all you knew, he was planning to blow up the vault.
question1d	You are approached by a frenzied vault scientist, who yells, "I'm going to put my quantum harmonizer in your photonic resonation chamber!" What's your response?	Say nothing, but slip away before the scientist can continue his rant.

Table K.1: Fallout study: conversation dialog

variable	conversation text	conversation option
question2a	While working as an intern in the clinic, a patient with a strange infection on his foot stumbles through the door. The infection is spreading at an alarming rate, but the doctor has stepped out for a while. What do you do?	Amputate the foot before the infection spreads.
question2b	While working as an intern in the clinic, a patient with a strange infection on his foot stumbles through the door. The infection is spreading at an alarming rate, but the doctor has stepped out for a while. What do you do?	Scream for help.
question2c	While working as an intern in the clinic, a patient with a strange infection on his foot stumbles through the door. The infection is spreading at an alarming rate, but the doctor has stepped out for a while. What do you do?	Medicate the infected area to the best of your abilities.
question2d	While working as an intern in the clinic, a patient with a strange infection on his foot stumbles through the door. The infection is spreading at an alarming rate, but the doctor has stepped out for a while. What do you do?	Restrain the patient, and merely observe as the infection spreads.
question4a	Congratulations! You made one of the Vault 101 baseball teams! Which position do you prefer?	Pitcher
question4b	Congratulations! You made one of the Vault 101 baseball teams! Which position do you prefer?	Catcher
question4c	Congratulations! You made one of the Vault 101 baseball teams! Which position do you prefer?	Designated Hitter
question4d	Congratulations! You made one of the Vault 101 baseball teams! Which position do you prefer?	None, you wish the vault had a soccer team

Table K.1: Fallout study: conversation dialog

variable	conversation text	conversation option
question6a	Old Mr. Abernathy has locked himself in his quarters again, and you've been ordered to get him out. How do you proceed?	Use a bobby pin to pick the lock on the door.
question6b	Old Mr. Abernathy has locked himself in his quarters again, and you've been ordered to get him out. How do you proceed?	Trade a vault hoodlum for his cherry bomb and blow open the lock.
question6c	Old Mr. Abernathy has locked himself in his quarters again, and you've been ordered to get him out. How do you proceed?	Go to the armory, retrieve a laser pistol, and blow the lock off.
question6d	Old Mr. Abernathy has locked himself in his quarters again, and you've been ordered to get him out. How do you proceed?	Just walk away and let the old coot rot.
question8a	A fellow Vault 101 resident is in possession of a Grognak the Barbarian comic book, issue no. 1. You want it. What's the best way to obtain it?	Trade the comic book for one of your own valuable possessions.
question8b	A fellow Vault 101 resident is in possession of a Grognak the Barbarian comic book, issue no. 1. You want it. What's the best way to obtain it?	Steal the comic book at gunpoint.
question8c	A fellow Vault 101 resident is in possession of a Grognak the Barbarian comic book, issue no. 1. You want it. What's the best way to obtain it?	Sneak into the resident's quarters, and steal the comic book from his desk.
question8d	A fellow Vault 101 resident is in possession of a Grognak the Barbarian comic book, issue no. 1. You want it. What's the best way to obtain it?	Slip some knock out drops into the resident's Nuka-Cola, and take the comic book when he's unconscious.

Appendix L

Fallout study descriptives

Table L.1: Fallout study game variables: descriptive statistics

Variable Name	N	Range	Min	Max	Mean	Std. Dev
Group 1						
Total pooled move	36	28	1	29	7.61	5.867
Total Pooled Con Game	36	16	13	29	21.08	3.996
Binnen de tijd klaar of niet	36	1	0	1	.11	.319
Group 2						
Pooled move phase 3	36	17	0	17	3.86	4.141
Pooled move phase 4	36	11	1	12	3.75	2.234
Pooled move area 1	36	10	0	10	2.28	2.514
Pooled move area 2	36	6	0	6	.89	1.410
Pooled move area 3	36	5	0	5	.69	1.191
Pooled move area 4	36	9	0	9	2.44	1.796
Pooled move area 5	36	3	0	3	1.31	.822
Group 3						
Total Pooled Con Ph3	36	10	8	18	13.69	2.528
Total Pooled Con Ph4	36	12	3	15	7.39	2.861
Group 4						
Pooled Amata Con	36	2	1	3	2.50	.697
Pooled Beatrice Con	36	1	1	2	1.92	.280
Pooled Butch Con	36	2	0	2	1.39	.728
Pooled Dad Con	36	4	2	6	3.22	.929
Pooled Jonas Con	36	2	0	2	1.22	.485
Pooled Old Lady Palmer Con	36	1	0	1	.75	.439
Pooled Overseer Con	36	1	0	1	.83	.378
Pooled Stanley Con	36	2	0	2	1.72	.701
Pooled Security Con	36	1	0	1	.14	.351
Pooled Amata 2 Con	36	4	0	4	.44	.909

Table L.1: Fallout study game variables: descriptive statistics

Variable Name	N	Range	Min	Max	Mean	Std. Dev
Pooled Butch 2 Con	36	3	0	3	.97	1.082
Pooled Dad 2 Con	36	9	0	9	3.61	1.975
Pooled Brotch Con	36	4	0	4	1.94	.715
Pooled Paul Con	36	1	0	1	.25	.439
Pooled Stanley 2 Con	36	0	0	0	.00	.000
Pooled Wally 2 Con	36	3	0	3	.17	.609
Group 5						
amac01	36	1	0	1	.11	.319
amac02	36	0	0	0	.00	.000
amac03	36	1	0	1	.33	.478
amac04	36	1	0	1	.06	.232
amac05	36	1	0	1	.03	.167
amac06	36	1	0	1	.03	.167
amac07	36	1	0	1	.03	.167
amac08	36	1	0	1	.08	.280
amac09	36	1	0	1	.03	.167
amac10	36	1	0	1	.06	.232
amac11	36	1	0	1	.22	.422
amac12	36	1	0	1	.67	.478
amac13	36	1	0	1	.11	.319
amac14	36	1	0	1	.33	.478
amac15	36	1	0	1	.28	.454
amac16	36	1	0	1	.14	.351
beac01	36	1	0	1	.81	.401
beac02	36	1	0	1	.08	.280
beac03	36	1	0	1	.03	.167
beac04	36	1	0	1	.53	.506
beac05	36	1	0	1	.28	.454
beac06	36	1	0	1	.19	.401
butc01	36	1	0	1	.36	.487
butc02	36	1	0	1	.17	.378
butc03	36	1	0	1	.08	.280
butc04	36	1	0	1	.11	.319
butc05	36	1	0	1	.50	.507
butc06	36	1	0	1	.17	.378
butc07	36	0	0	0	.00	.000
butc08	36	0	0	0	.00	.000
dadc01	36	1	0	1	.31	.467
dadc02	36	0	0	0	.00	.000
dadc03	36	1	0	1	.03	.167
dadc04	36	1	0	1	.50	.507
dadc05	36	1	0	1	.50	.507

Table L.1: Fallout study game variables: descriptive statistics

Variable Name	N	Range	Min	Max	Mean	Std. Dev
dadc06	36	0	0	0	.00	.000
dadc07	36	1	0	1	.03	.167
dadc08	36	1	0	1	.08	.280
dadc09	36	0	0	0	.00	.000
dadc10	36	1	0	1	.08	.280
dadc11	36	0	0	0	.00	.000
dadc12	36	1	0	1	.97	.167
dadc13	36	1	0	1	.03	.167
dadc14	36	1	0	1	.44	.504
dadc15	36	1	0	1	.25	.439
dadc16	36	0	0	0	.00	.000
jonc01	36	1	0	1	.31	.467
jonc02	36	1	0	1	.11	.319
jonc03	36	1	0	1	.53	.506
jonc04	36	1	0	1	.17	.378
jonc05	36	1	0	1	.03	.167
jonc06	36	1	0	1	.08	.280
olpc01	36	1	0	1	.17	.378
olpc02	36	1	0	1	.50	.507
olpc03	36	1	0	1	.08	.280
ovec01	36	1	0	1	.14	.351
ovec02	36	1	0	1	.69	.467
stac01	36	1	0	1	.69	.467
stac02	36	1	0	1	.06	.232
stac03	36	1	0	1	.08	.280
stac04	36	1	0	1	.03	.167
stac05	36	1	0	1	.72	.454
stac06	36	1	0	1	.03	.167
stac07	36	1	0	1	.11	.319
secc01	36	1	0	1	.06	.232
secc02	36	1	0	1	.08	.280
secc03	36	0	0	0	.00	.000
amac17	36	1	0	1	.17	.378
amac18	36	1	0	1	.06	.232
amac19	36	2	0	2	.08	.368
amac20	36	1	0	1	.14	.351
butc09	36	0	0	0	.00	.000
butc10	36	1	0	1	.06	.232
butc11	36	1	0	1	.03	.167
butc12	36	1	0	1	.31	.467
butc13	36	1	0	1	.06	.232
butc14	36	1	0	1	.03	.167

Table L.1: Fallout study game variables: descriptive statistics

Variable Name	N	Range	Min	Max	Mean	Std. Dev
butc15	36	1	0	1	.39	.494
butc16	36	1	0	1	.06	.232
butc17	36	1	0	1	.06	.232
dadc17	36	1	0	1	.86	.351
dadc18	36	1	0	1	.36	.487
dadc19	36	4	0	4	1.08	.732
dadc20	36	3	0	3	.64	.683
dadc21	36	1	0	1	.58	.500
dadc22	36	1	0	1	.08	.280
broc01	36	1	0	1	.25	.439
broc02	36	1	0	1	.33	.478
broc03	36	1	0	1	.19	.401
broc04	36	1	0	1	.03	.167
broc05	36	1	0	1	.19	.401
broc06	36	1	0	1	.72	.454
broc07	36	1	0	1	.11	.319
broc08	36	1	0	1	.03	.167
broc09	36	1	0	1	.06	.232
broc10	36	0	0	0	.00	.000
broc11	36	1	0	1	.03	.167
broc12	36	0	0	0	.00	.000
broc13	36	0	0	0	.00	.000
pauc01	36	1	0	1	.19	.401
pauc02	36	1	0	1	.06	.232
stac08	36	0	0	0	.00	.000
stac09	36	0	0	0	.00	.000
walc01	36	1	0	1	.03	.167
walc02	36	1	0	1	.06	.232
walc03	36	1	0	1	.08	.280
walc04	36	0	0	0	.00	.000
question1a	36	1	0	1	.22	.422
question1b	36	1	0	1	.25	.439
question1c	36	1	0	1	.22	.422
question1d	36	1	0	1	.08	.280
question2a	36	1	0	1	.19	.401
question2b	36	1	0	1	.03	.167
question2c	36	1	0	1	.53	.506
question2d	36	1	0	1	.03	.167
question3a	36	1	0	1	.14	.351
question3b	36	1	0	1	.11	.319
question3c	36	0	0	0	.00	.000
question3d	36	1	0	1	.53	.506

Table L.1: Fallout study game variables: descriptive statistics

Variable Name	N	Range	Min	Max	Mean	Std. Dev
question4a	36	1	0	1	.25	.439
question4b	36	1	0	1	.06	.232
question4c	36	1	0	1	.14	.351
question4d	36	1	0	1	.33	.478
question5a	36	1	0	1	.11	.319
question5b	36	1	0	1	.19	.401
question5c	36	1	0	1	.11	.319
question5d	36	1	0	1	.36	.487
question6a	36	1	0	1	.42	.500
question6b	36	0	0	0	.00	.000
question6c	36	1	0	1	.33	.478
question6d	36	1	0	1	.03	.167
question7a	36	1	0	1	.06	.232
question7b	36	1	0	1	.31	.467
question7c	36	0	0	0	.00	.000
question7d	36	1	0	1	.42	.500
question8a	36	1	0	1	.53	.506
question8b	36	1	0	1	.06	.232
question8c	36	1	0	1	.14	.351
question8d	36	1	0	1	.06	.232
question9a	36	1	0	1	.28	.454
question9b	36	1	0	1	.17	.378
question9c	36	1	0	1	.03	.167
question9d	36	1	0	1	.31	.467
question10	36	1	0	1	.78	.422
Trigger (10j) bar	36	4	0	4	.53	1.000
Trigger (10j) gesloten deur clinic	36	1	0	1	.11	.319
Trigger (10j) gesloten deur bovenin	36	2	0	2	.19	.577
Trigger (10j) gesloten deur appartements	36	2	0	2	.19	.467
Trigger (10j) gesloten deur reactor	36	2	0	2	.11	.398
Trigger (10j) gesloten deur filter room	36	2	0	2	.36	.593
Trigger (10j) deur bar	36	8	0	8	1.75	1.779
Trigger (10j) voor doorgang feestje	36	3	0	3	1.25	.996
Trigger (10j) kruispunt	36	6	0	6	1.19	1.142
Trigger (10j) tweede kruispunt	36	3	0	3	.92	.649
Trigger (10j) kamer Jonas	36	2	0	2	.39	.549
Trigger (16j) gesloten deur	36	3	0	3	.39	.728
Trigger (16j) gesloten deur atrium	36	3	0	3	.22	.637
Trigger (16j) achter bureau dad	36	2	0	2	.25	.554
Trigger (16j) achter scherm	36	1	0	1	.08	.280
Trigger (16j) voor Stanley	36	3	0	3	.58	.770
Fought tunnelsnakes	36	1	0	1	.42	.500

Table L.1: Fallout study game variables: descriptive statistics

Variable Name	N	Range	Min	Max	Mean	Std. Dev
Participated in the GOAT	36	10	0	10	7.78	4.216

Appendix M

Fallout study significant correlations

Table M.1 contains all the correlations found during the one-tailed correlation test which was performed as part of the linear regression analysis for the FALLOUT 3 experiment in Chapter 7. Correlations with the * symbol are significant at the $p < 0.05$ level, correlations with the ** symbol are significant at the $p < 0.01$ level. Correlations without * or ** are included for the sake of completeness. The following abbreviations apply: O = Openness, C = Conscientiousness, E = Extraversion, A = Agreeableness, and N = Neuroticism.

Table M.1: FALLOUT study: correlations

	O	C	E	A	N
	Group 1				
Total Pooled Con Game	-.087	-.116	.067	-.052	-.301*
Total pooled move	.260	-.062	.206	-.432**	.058
Finished Game	-.044	.239	.069	.030	-.379*
	Group 2				
Pooled move phase 1	.246	-.013	.149	-.482**	.109
Pooled move area 1	.248	.020	.008	-.449**	.179
Pooled move area 2	.205	.054	.211	-.420**	.113
	Group 3				
Total Pooled Con Ph1	-.220	.120	.142	.039	-.369*
	Group 4				
Pooled Amata Con	-.064	.200	.077	.021	-.365*
Pooled Jonas Con	-.164	-.200	.319*	-.010	-.261
Pooled Overseer Con	.087	.155	.007	.341*	-.410**
Pooled Amata 2 Con	-.017	-.356*	.021	-.042	-.124
Pooled Brotch Con	-.160	.001	-.091	-.134	-.366*
Pooled Wally 2 Con	.415**	-.066	.046	-.212	-.034
	Group 5				

Table M.1: FALLOUT study: correlations

	O	C	E	A	N
amac01	-.293*	.052	.123	.105	-.501**
amac05	-.086	.220	.288*	.158	-.056
amac07	-.301*	-.337*	.196	-.014	-.056
amac08	-.027	.061	.023	-.281*	.089
amac09	.021	-.337*	.013	-.186	-.056
amac10	-.047	-.164	-.311*	-.267	.071
amac12	-.087	.246	.266	.270	-.318*
amac13	-.068	-.356*	-.117	-.300*	.214
amac14	.162	.336*	.021	.150	-.049
amac15	-.081	-.093	-.021	.011	-.397**
beac05	-.199	.152	.316*	.042	-.088
beac06	.238	-.101	-.306*	-.149	-.074
butc01	-.530**	.177	.088	.025	-.175
butc02	.386*	-.008	.115	.000	.085
butc03	-.281*	-.006	.350*	.077	-.036
butc04	-.012	-.123	.027	-.345*	.049
dadc03	-.086	-.003	.288*	.244	-.161
dadc04	.088	.018	.105	.028	.467**
dadc05	-.088	-.018	-.105	-.028	-.467**
dadc07	-.301*	-.337*	.196	-.014	-.056
dadc08	-.218	.061	.295*	.026	.026
dadc10	-.218	.061	.295*	.026	.026
dadc12	.086	.226	.354*	.186	-.155
dadc13	-.086	-.226	-.354*	-.186	.155
dadc15	.356*	.201	-.305*	-.245	.090
jonc04	-.134	-.254	-.007	-.114	-.286*
jonc05	.128	-.003	.288*	-.100	-.371*
jonc06	-.090	.061	.295*	.077	.151
olpc03	.292*	-.138	-.086	.077	-.036
ovec01	-.205	-.166	-.144	-.157	-.282*
ovec02	.224	.250	.114	.394**	-.120
stac03	.228	-.138	-.468**	-.230	.214
stac06	.128	-.003	.288*	-.100	-.371*
stac07	.100	-.239	-.165	-.345*	.049
secc01	-.124	.315*	.018	-.021	-.004
amac17	.102	-.303*	.034	.000	-.425**
amac20	.049	-.378*	-.013	-.157	-.132
butc11	.128	-.226	.013	-.014	-.371*
butc13	-.047	-.004	.150	.103	-.381*
dadc22	-.154	-.204	-.359*	-.026	.214
broc03	.061	-.009	-.077	.280*	-.205
broc09	.030	-.324*	-.113	-.082	-.080

Table M.1: FALLOUT study: correlations

	O	C	E	A	N
pauc02	-.278	.075	.413**	-.021	-.004
walc01	.342*	.108	.104	-.272	.155
walc02	.337*	-.084	.084	-.206	-.155
walc03	.419**	-.138	-.032	-.128	-.036
question1c	-.104	-.098	.113	.125	-.342*
question1d	.292*	.061	-.032	.0260	.214
question2a	.372*	-.055	.075	-.363*	-.292*
question2b	-.301*	-.337*	.196	-.014	-.056
question2d	.021	.108	.013	.330*	-.056
question3d	-.293*	.017	-.041	-.061	.068
question4a	-.051	.327*	.044	-.114	.090
question4b	-.278	-.404**	.018	-.082	.147
question4d	.087	-.168	.085	.150	-.379*
question6c	-.212	.181	.341*	.330*	-.196
question8c	-.052	-.166	.379*	.170	-.532**
question9c	-.086	.220	.288*	.158	-.056
Trigger (10j) bar	.262	.027	.010	-.361*	.157
Trigger (10j) gesloten deur bovenin	.290*	.090	.238	-.327*	-.052
Trigger (10j) gesloten deur appartements	.128	.111	.294*	-.251	.011
Trigger (10j) deur bar	.203	.013	.006	-.431**	.165
Trigger (10j) doorgang feestje	-.022	-.116	.387**	-.022	-.189
Trigger (10j) kruispunt	.303*	-.036	.080	-.278	.005
Trigger (16j) gesloten deur	-.007	.016	-.085	-.282*	.183
Trigger (16j) gesloten deur atrium	.016	-.065	.242	-.165	-.446**
Trigger (16j) achter scherm	.292*	.061	.186	-.332*	-.162
Trigger (16j) voor Stanley	.442**	-.062	-.022	-.438**	-.025

Appendix N

Interview validation of game personality

Below we provide some ideas on interview validation of games personality of which we came across during research. However, because there are several methodological difficulties with this research, we decided not to break the stream of information for the main chapters. We decided to publish the information obtained during this investigation as an appendix.

The *lexicographic method* (described in more detail in Subsection 2.5.2) has been quite successful in providing personality classifications. The lexicographic method has some difficulty in providing explanations for behaviour on a small scale, referred to in the literature as *actual behaviour*. In contrast to self-reporting on behaviour, we aim to investigate actually observable behaviour (shortened to: *actual behaviour*). The prediction of *actual behaviour* often has a low reliability and the obtained results are hard to replicate (Back and Egloff, 2009).

The difficulties in predicting *actual behaviour* can be found in many cases. Human behaviour, outside the laboratory, is influenced by many variables (for interpretation, see Cohen in Section 4.4). In the past, the solution for reducing this unwanted variation was often measuring behaviour in many different situations and circumstances, thereby averaging out the unwanted variation. Gathering data this way was an attempt to average out the effects of the various confounding factors. The disadvantage of this approach is that personality scores become less specific and, more importantly, less tailored to the individual and more tailored to interpretations based on averages of behaviour across a population.

Recently, there have been requests for a more behaviouralistically oriented approach to personality. The consensus has been reached that advances in technology, in particular ubiquitous data collection and advanced data analysis, have provided us with new and accurate ways in which to perform personality profiling by measuring *actual behaviour*(Furr, 2009).

N.1 The game: the poisoned lake (continued)

In this appendix, we use the same game that we have used in Chapter 5. The only difference with Chapter 5 is that, in this chapter, the instructions for controlling the game were given verbally instead of in an instruction booklet.

N.2 Experimental setup

Below we present the investigation into the effectiveness of personality tests and gameplay data on predicting real world behaviour. We provide details on the participants, the experimental procedure, and the measures we have used during the investigation. Thereafter, we discuss the questionnaire, an interview conducted by (Back and Egloff, 2009) and behavioural criteria.

N.2.1 Participants

Our participants were 37 students of Tilburg University, the Netherlands. There were 22 male and 15 female participants. The students were recruited from the student population. Participants received research credits for participating in the study.

N.2.2 Procedure

The investigation consisted of two parts: (1) an online NEO-PI-R personality questionnaire, and (2) a laboratory session containing an interview and the playing of a game. The game we used is the same game and module discussed in Chapter 5. Students were able to start the online questionnaire as soon as they had registered for the experiment. An appointment for the laboratory session was scheduled within a month of completing the online survey.

In the laboratory session, after signing a consent form, participants were told that the session would contain a short interview and the playing of a video game. A video recording was started and participants were seated opposite to the interviewer. Participants faced the video camera. What followed was an interview which lasted for 3 minutes on average (between 2 and 4 minutes). The interview contained a selection of the personality measures used by Back and Egloff (2009). The sections that were selected can be found in Appendix O. Participants were free to answer as they liked and to take as much time as they required.

After the interview was finished, the video recording was stopped and the participants were seated in front of a computer with a mouse and keyboard. Participants received a brief instruction about the controls of the game and were told that they could play the game until they completed it or until sixty minutes had passed. All participants were able to complete the game within the time limit.

Participants required 30 to 45 minutes to complete the online questionnaire and 45 to 65 minutes to complete the laboratory session, in which the interview was conducted and a game was played. There were no significant time differences between male and female participants. Gameplay lasted between 20 and 60 minutes (40 minutes on average) and every participant was able to complete the game. The average gameplay duration was slightly shorter than in the Chapter 5 experiment. One possible reason for the decreased time may be the addition of a short verbal explanation of the controls at the start of the gaming part of the lab session.

N.2.3 Measures

There were three parts of the data collection in this investigation, (1) the online questionnaire, (2) the interview containing several behavioural criteria, and (3) the computer game.

Questionnaire

An online version of the NEO-PI-R questionnaire was created. The NEO-PI-R is described in Subsection 2.5.3. The online version of the questionnaire contains the same selection of questionnaire items in the same order as the original version. The response options are also the same as in the original version. So this measure was the same as the measure used in Chapter 5.

Interview

In our experiment we used an interview that was adapted from Back and Egloff (2009). A protocol was written for the interview. It can be found in Table N.1. The interview was chosen based on literature by Back and Egloff (2009). The interview consisted of three parts (1) Small talk (ST), (2) self-introduction (SI), and (3) a vision on the future (VF). For each of the three parts the interviewer followed an a priori defined script for the conversation.

During ST the interviewer asked the participant to say their name for the video. The participant was asked which study he followed and whether he enjoyed the study. After this, the participant was asked if their telephone was set to off or silent mode. After ST followed SI. Here participants were asked to "give a description of themselves, what their hobbies were and in what way they filled the spare time they had next to their studies. Finally, in the VF part, the participant was asked to give a prediction of what he thought the future would look like in 15 years. The participant was asked to describe the world in general as well as his own life.

The interviewer did not interrupt the stories told by the participant except for giving small encouraging signs (like "uhuh" etc.) and an occasional short summary (maximum one sentence) to indicate comprehension to the participant. Sometimes it was necessary to remind the participant of the question. An example would be in VF if a description of the world was given but a description of the participant himself was omitted.

The purpose of the interview was to provide the participant with opportunities to express himself and display personality-related behaviour. The interview questions were meant to evoke responses in which the participant was not restrained in the way he answered.

Behavioural criteria

The behavioural criteria for the interview were created a priori. They were chosen from the measures used in Back and Egloff (2009). Back and Egloff (2009) created over 50 criteria in order to predict the "Big-Five" personality traits. This approach resulted in ten criteria per trait on average. However, unlike Back and Egloff (2009), we reduced the video duration by

Table N.1: Video interview protocol.

Reception
Participant enters. Shake hand, introduce yourself as the experimentor, thank participant for participating. Have participant fill out consent form without the video camera activated. Explain that the video interview will commence. Activate the camera and sit down. State the date and time for the video.
Small talk
"Please state your name." "What is your major?" "How do you feel about studying?" "Is your mobile phone off?"
Self introduction
"Could you tell me a bit about yourself, about your hobbies and interests and so on?"
Vision of the future
"Please describe what you think your life and the world in general will look like in 15 years."
Continue with the game
Finish the interview conversation, state that we are now proceeding with the next part of the experiment and turn off the video camera. Relocate the participant to the desk with the computer and activate the game.

using only the ST, SI, and VF situations and omitting the other situations. This reduced our video duration to a maximum of 15 minutes while preserving 36 of the original 53 criteria. Back et al. performed an extensive validation by requiring experts to examine the appropriateness of each criterium for its given trait. The selection of the behavioural criteria we used is given in Appendix O.

N.3 Results

Before we could start our analyses we needed to compute several variables based on the procedures described in Chapter 5 (pooling variables) and based on Back and Egloff (2009) (aggregating interview variables). After computing the variables, we conducted the analysis on the NEO-PI-R, the game, and the interview.

In Subsection N.3.1 we explain our computations. In Subsection N.3.2 we provide an overview of our data by presenting the descriptives. In Subsection N.3.3 we present the results of our correlation analysis. In Subsection N.3.4 we present the results of our combined dataset analysis. In Subsection N.3.5 we show the reliability analysis results.

N.3.1 Variable computation

The first step of the analysis was creating the pooled variables for the groups described in Chapter 5. The second step was transforming the responses in the interview into standardised z-scores. Following this standardisation, trait scores were computed by taking the mean of the standardised scores for the interview items of each trait. The aggregated interview data produced trait measures for each of the Big Five traits.

N.3.2 Descriptives

The z-scores resulting from the aggregation procedure by Back and Egloff (2009) can be found in Table N.2. The table shows that there is considerable variance in the scores for the trait openness compared to the other traits. The other traits seem to have a roughly equal range of scores.

Table N.2: Back and Egloff (2009) interview descriptives.

Trait	Minimum	Maximum	M	SD
Openness	-1.15	2.76	.000	.76116
Conscientiousness	-1.01	1.43	.000	.53563
Extraversion	-.99	1.20	.000	.54568
Agreeableness	-.95	1.01	.000	.43900
Neuroticism	-.93	1.27	.000	.56469

Via the creation of the pooled variables and the aggregation of the interview scores, we compared the descriptives of the NEO-PI-R scores, subsequently we generated and examined the scores. Table N.3 shows the descriptive statistics of the NEO-PI-R questionnaire. The scores in Table N.2 are not in the same range as the scores in Table N.3. The reason for this difference is that, unlike the interview scores, the NEO-PI-R scores originate from a normalised method that has been tested on a large population.

Table N.3: NEO-PI-R trait descriptives.

Trait	Minimum	Maximum	M	SD
Openness	1	9	5.58	1,588
Conscientiousness	1	8	4.21	2.002
Extraversion	2	9	6.24	1.807
Agreeableness	1	7	3.66	1.921
Neuroticism	1	9	5.34	1.977

Our descriptives show that the trait scores in the interview as well as in the NEO-PI-R questionnaire follow the normal distribution.

N.3.3 Correlation analysis

Table N.4 shows the correlation between the traits found using the NEO-PI-R questionnaire and the interview aggregated trait scores. All correlations have $p > .05$ and are therefore not statistically significant. The first column lists the OCEAN trait aggregates found in the interview. The top row lists the OCEAN traits found by the NEO-PI-R questionnaire.

Table N.4: Correlations between the NEO-PI-R trait scores and the interview trait scores.

	O	C	E	A	N
Openness	.083	-.259	.152	-.136	-.004
Conscientiousness	.044	.135	.062	.306	-.201
Extraversion	.156	-.084	.168	-.233	.209
Agreeableness	.078	-.070	-.162	.161	.079
Neuroticism	.005	.086	-.302	-.059	.212

All correlations have $p > .05$.

Back and Egloff (2009) found correlations displayed in Table N.5 with significance levels of $p < .001$. Our results were not significant and the correlation levels we found were almost half those found in the experiment of Back and Egloff (2009).

Table N.5: NEO-PI-R - Interview correlations.

Trait	r (Back and Egloff (2009))	r (found in this experiment)
Openness	.310**	.083
Conscientiousness	.300**	.135
Extraversion	.380**	.168
Agreeableness	.350**	.161
Neuroticism	.360**	.212

** $p < .01$. * $p < .05$ (not in table).

N.3.4 Combined datasets

The results in Table N.6 show that our a game model can be created to predict scores on the interview aggregates. These scores were predicted with the regression analysis of the game data on the NEO-PI-R trait scores.

These results show that the combined dataset is more effective at predicting agreeableness scores.

Correlation

The correlation studies between the aggregated interview personality traits and the game personality scores in this chapter show small effect sizes and no correlation effects for all of

Table N.6: Linear regression analysis for game variables on interview FFM scores.

Trait	R^2 (Interview)	R^2 (Game)
Openness	1.00	.768
Conscientiousness	1.000	.559
Extraversion	.910	.351
Agreeableness	.548	.724
Neuroticism	1.000	.568

the personality traits (see Table N.7, column: Interview). One of the possible causes is that the linear regression model that produces the game personality scores is trained on too few participants. Training a model on too few cases might cause overfitting (Hawkins, 2004). In order to reduce the chances of overfitting more participants could be used to train the model on. We refer to in Chapters 4, 5, and 7, for more information about the interpretation of correlation analyses.

The research in this chapter is similar to the research in Chapter 5. We use the same game and the same personality test in both chapters. Therefore, it is possible to combine the data for both chapters and train a regression model on that data set. Because the total number of participants nearly doubles, the model is much more reliable. We can then try to correlate the scores generated by the new regression model to the interview personality scores in this chapter.

In Table N.7 the correlations between (1) the game, (2) the NEO-PI-R, and (3) the interview trait scores for three different models can be found as follows. (Ad1) A multiple linear regression model trained on the Chapter 5 dataset, (Ad2) A multiple linear regression model trained on the Chapter 6 dataset, and (Ad3) A multiple linear regression model trained on the Chapter 5 and Chapter 6 datasets combined. The correlation analysis between game and NEO-PI-R data was performed for the full test subject set in both chapters; the combined number reads $N = 80$. The correlation analysis between the game and the interview could only be done for the subset of the participants that participated in the experiments for this chapter ($N = 37$). This may be the cause of lower reliability for the analysis. The correlation for the game model based on Chapter 6 is a repeat of the analysis found in Table N.4.

Table N.7: Correlational between game, NEO-PI-R, and interview scores.

Game	NEO-PI-R	Interview
Openness	.876**	.010
Conscientiousness	.748**	.013
Extraversion	.593**	-.049
Agreeableness	.851**	.227
Neuroticism	.753**	.233

** $p < .01$. * $p < .05$ (not in table).

N.3.5 Reliabilities

In order to test whether the interview was as reliable as the interview in Back and Egloff (2009) we calculated Cronbach's α for every set of test items. The results can be found in Table N.8. The reliabilities were lower than those found in Back's original study. An acceptable value for judging reliability as "good enough" is a value of .7 or higher (Kline, 1999).

Table N.8: Interview item reliabilies found for each trait.

Trait	α	α (Back article)
Openness	.54	.73
Conscientiousness	.31	.85
Extraversion	.07	.87
Agreeableness	.07	.75
Neuroticism	.55	.84

The reliability analysis shows that for three of the traits (openness, extraversion, and neuroticism) reliability could be improved by deleting an item from the scale. Reliability for openness increases to .721 by deleting one item. Reliability for extraversion can be raised to .549 by removing one item from the item-set and agreeableness can be raised to .255. However, the other two traits do not provide good candidate items in order to increase reliability.

N.4 Discussion

In this appendix we compared the NEO-PI-R personality questionnaire to the NWN game data with regards to predicting real world behaviour. The real world behaviour was in the form of a structured interview adapted from Back and Egloff (2009).

We used only 36 of Back's original 50 behavioural criteria in order to reduce the amount of time required to perform the interview. As can be seen in Table N.6, our linear regression analysis was able to fit a model to predict the aggregated OCEAN scores of the interview using our game data as predictors.

We also performed 9 correlation analyses in order to investigate the ability of a game-based personality model to predict NEO-PI-R scores and in order to predict aggregated interview scores. We performed the first correlation analysis to see whether the linear regression model trained on the dataset of this appendix correlated to the aggregated interview scores. We performed three correlation analyses of regression models for the interrelations between datasets for Chapters 5, 6, and the combined data for both chapters.

Although we may state that we have somewhat successfully created a model that can predict the scores in our interview we are facing an unexpected result. In the original article by Back and Egloff (2009) a considerable reliability between interview items was shown. Based on this reliability, the subset of items that we have selected for our research should have (according to the reliabilities found by Back and Egloff (2009)) have remained reliable in measuring personality traits. However, no relationship between the interview scores and the personality scores of the NEO-PI-R personality test was found (see Table N.5). We present four possible explanations for the absence of the expected effects.

1. Back and Egloff (2009) used 130 participants in their research whereas we only used 37. It is possible that the amount of subjects we have used is too small to coalesce into statistically significant results.
2. The interviewer might have performed the interview differently from the way the original interviewers did their experiment (cf. Back and Egloff, 2009), resulting in a significantly different result. Although the interview items in the Back and Egloff (2009) article have been extensively described, there was still room for interpretation. Back and Egloff (2009) did neither provide a clear explanation on the location of the interviewer and the location of the camera during their experiment, nor did they provide a clear explanation of their laboratory setup.
3. The experiment described in this appendix was rated by only one rater. Multiple raters in a double-blind setup may provide other results (preferably more convincingly).
4. Back and Egloff (2009) used personality experts to construct their behavioural criteria. When our raters examined the items and consulted experts in our environment there was disagreement on which personality trait should be measured by several of the items. This ambiguity of items should not be present according to the reliability measured by Back and Egloff (2009).

Future work may benefit from double-blind testing and multiple raters. Additional interviewers and a different set of behavioural criteria also deserve consideration. A repeat of the original experiment by Back and Egloff (2009) is also advised because replication of the original results was not successful.

N.5 Appendix conclusions

In this appendix we have investigated the relationships between game behaviour, questionnaire scores, and *actual behaviour* (in the form of interview responses). We attempted to answer how game-based personality tests compare to personality questionnaires in predicting real-world behaviour.

From the results, we may conclude that we can predict some of the real world behaviour using game data by creating a linear model using regression analysis provided that the interviews by Back and Egloff (2009) are an adequate yardstick for *actual behaviour*. However, as stated above, we could not reproduce the effects found by Back and Egloff (2009). So, we may question either our researched approach or the results found by Back and Egloff (2009), or both. With respect to the answer, we admit that our game-based personality models fare poorly when trying to predict interview behaviour.

Appendix O

Behavioural criteria

In Table O.1 we provide an overview of the behavioural criteria. Table O.1 contains three columns per trait. The first column displays the name of the variable. The second column provides a short description of the variable. The third column shows the type of the variable. Rating means the variable is rated by a trained rater (RVO). The rater observes the recorded behaviour and provides an appropriate rating. Count means the variable was an objective measure represented by a word or behaviour count. Neuroticism of speech is the number of filler words used in the search, examples are sounds like “uhh” or “ehm” while the participant pauses his speech. The variable type “LIWC” refers to the linguistic enquiry and word count by Pennebaker et al. (2007). Openness 3 refers to the amount of personal information the participant presents by himself (without prompts from the interviewer) during ST.

Table O.1: Behavioural criteria selected from the Back and Egloff (2009) experiment.

Variable name	Behavioural criteria	Variable type
	Openness	
Openness 1	Global transcript SI	Rating
Openness 2	Verbal eloquence SI	Rating
Openness 3	Open answers in small-talk situation	Rating
	Conscientiousness	
Conscientiousness 1	Global transcript SI	Rating
Conscientiousness 2	Minutes too late in attending experiment	Count
Conscientiousness 3	Understandability in ST	Rating
Conscientiousness 4	Slouching body posture SI	Rating
Conscientiousness 5	Formal dress	Rating
	Extraversion	
Extraversion 1	Global behavior SI	Rating
Extraversion 2	Global behavior vision of the future	Rating
Extraversion 3	Global transcript SI	Rating
Extraversion 4	Expressivity of facial expression SI	Rating
Extraversion 5	Loudness of voice SI	Rating
Extraversion 6	Number of words SI	LIWC
Extraversion 7	Own questions during small talk	Count
Extraversion 8	Second-person pronouns SI	LIWC
Extraversion 9	Other references SI	LIWC
Extraversion 10	Stylish dress	Rating
Extraversion 11	Flashy dress	Rating
	Agreeableness	
Agreeableness 1	Global transcript SI	Rating
Agreeableness 2	Friendly voice in small talk	Rating
Agreeableness 3	Attentive body posture in small talk	Rating
Agreeableness 4	Number of swear words SI	LIWC
Agreeableness 5	Relative frequency of other- versus self-words SI	LIWC
Agreeableness 6	Words related to social processes SI	LIWC
Agreeableness 7	Words related to family SI	LIWC
	Neuroticism	
Neuroticism 1	Global behavior SI	Rating
Neuroticism 2	Global behavior vision of the future	Rating
Neuroticism 3	Global transcript SI	Rating
Neuroticism 4	Gaze aversion SI	Count
Neuroticism 5	Tense body posture SI	Rating
Neuroticism 6	Tense leg posture SI	Rating
Neuroticism 7	Silence during SI	Count
Neuroticism 8	Reassuring whether cell phone is switched off	(yes/no)
Neuroticism 9	Dysfluency of speech SI	LIWC
Neuroticism 10	Negations SI	LIWC
Neuroticism 11	Words related to anxiety and depression SI	LIWC

Summary

Computer games have existed for over 60 years and they are a popular medium of entertainment. Recently, computer games are being explored for other purposes such as education and assessment. Obviously, players of computer games vary in personality. We see these differences by looking at differences in play and by looking at the emotional and cognitive responses of players.

We aim to investigate the interaction between a player's psychology and the game content. More knowledge about this topic is desireable for two reasons: (1) to increase our control and understanding of the experience a player has while he is playing a game and (2) to be able to adapt content to suit a player. Currently, our knowledge of the psychology at work during game-play is limited and exploration of player psychology is our main interest.

The problem statement of this thesis is: *To what extent are games an appropriate means for measuring the differences between individuals based on psychological theories?* To investigate the problem statement we examine incongruity theory and personality theory. We investigate the influence of these two theories on expressed emotions, on behaviour in games, and on responses on personality tests. We conduct five investigations to explore the extent to which psychological theories can explain the differences between individuals.

In Chapter 2 we present the background on the psychological concepts connected to our research and to previous research of related topics. We extensively discuss the concept of modelling from different theoretical perspectives.

After presenting the background information, Chapter 3 starts by presenting our research on RQ1: *To what extent are games suitable for measuring incongruity?* Incongruity theory states that players should feel boredom in easy games, pleasure in balanced games, and frustration in hard games. We investigate the relationship between the level of complexity of the player's mental model and the emotions the same player expresses. We implement a game called GLOVE in which players can be confronted with a scenario of low, balanced, or high complexity (i.e., an easy, a balanced, or a hard game). From the results in Chapter 3 we may conclude that players feel frustrated when playing a hard game and that they feel pleasure when playing a balanced game. We are unable to conclude that players feel boredom when playing an easy game.

In Chapter 4 we investigate RQ2: *To what extent can games be used to measure complex psychological processes such as extraversion?* We investigate the process of extraversion by incorporating several extraversion experiments done in the past into a scenario programmed for the game NEVERWINTER NIGHTS. In this scenario, players follow a short storyline while they perform in-game tasks imported from extraversion literature. While the players play,

we record their responses on the in-game tasks. From our results we may conclude that, for 12 of the 21 recorded behaviours, correlations to extraversion or to one of the facets of extraversion are found. It is our opinion that, with additional fine-tuning, this approach can be used to measure extraversion to a larger extent. Extraversion is only one of the five personality traits in the "Big-Five" personality model.

In Chapter 5 we attempt to expand our investigation to all personality traits. There we investigate RQ3: *To what extent can a data-driven personality profile be created based on game behaviour?* We investigate all five traits of the "Big-Five" personality model. We wish to focus our behavioural measures on in-game behaviour, rather than on explicitly formulated replications of experiments. We consider a behavioural measure approach to be applicable in more game situations. We implement a new scenario for the game Neverwinter Nights. This scenario conforms to the scenarios found in commercial computer games. We attempt to provide the players with a broad range of possible responses to the situations encountered in the game in order to enable the free expression of the player's personality. Because this approach is data-driven we construct 217 unpooled game variables that record the player's movement, conversation, and general data in the game. We also formulate 43 pooled variables in order to investigate some of our a priori assumptions. We conduct correlation analyses for all game variables with the five personality traits. We also perform linear regression analyses. From our results we may conclude that personality effects for all five traits are expressed in game behaviour in both our correlation analyses and our regression analyses. Therefore, we are able to form a personality profile based on game behaviour. An approach with so many variables for these analyses runs the risk of overfitting for the number of experimental participants we used. In order to reduce the risk of overfitting, in the next chapter we focus on a theory-driven approach.

In Chapter 6 we investigate RQ4: *To what extent does a theory-driven model explain personality in games?* In order to investigate this question we formulate eleven behavioural criteria based on behavioural descriptions given in the "Big-Five" literature. We compute the criteria from the dataset gathered in the investigation of Chapter 5. We conducted a correlations analysis between the eleven theoretical variables and the "Big-Five" personality traits. From the results we may conclude that the neuroticism and the agreeableness variables lead to correlation with their respective personality traits. One additional conscientiousness variable correlated with the neuroticism trait. This approach shows us that the assumptions we made regarding what behaviour to expect to occur for each personality trait should be carefully tested. In the next chapter we focus on validating the results we found in this chapter and in Chapter 5.

In Chapter 7 we investigate RQ5: *To what extent can our models of personality in games be validated in different games?* In order to see if the results we found in the previous chapter are valid across games we implement the game behaviour variables from Chapter 5 in the starting scenario of the commercially successful game Fallout 3. After collecting the data, we analyse the dataset using both the data-driven as well as the theory-driven analyses. We may conclude that the results from the data-driven approach can be replicated to some extent but the results from the theory-driven approach are replicated less successfully.

In Chapter 8 we present a discussion on the insights that were gained during the investigations in this thesis. We may conclude that validation, by a sufficient number of participants, and control variables are points of attention when performing psychological research in a game setting. We also discuss the measure of success that we have had with our chosen approaches. We attempt to provide advice in order to avoid pitfalls and we attempt to illuminate areas of interest that may merit future research attention, such as skill, preference, intelligence, and demographics.

In Chapter 9 we present our conclusions. We provide our answers to the research questions posed in Chapter 1 and we provide our answer to the problem statement posed in that same chapter. Finally, we may conclude that games can successfully be used to quantify individual player differences based on psychological theories but that care must be taken with regards to analysis, validation, the number of participants used, and the interpretation of results.

Samenvatting

Computerspelen bestaan nu al meer dan 60 jaar en zijn een populair medium voor entertainment. Recentelijk is er een groeiende interesse ontstaan voor computerspelen met educatieve- en toetsingsdoeleinden. De spelers van computerspelen variëren aanzienlijk betreffende hun voorkeuren, cognities en vaardigheden. Ons onderzoek is er op gericht deze verschillen te onderzoeken door, tijdens of na het spelen, te kijken naar de emotionele reacties en het gedrag van de spelers.

Wij onderzoeken de interactie tussen de psychologie van een speler en de inhoud van een spel. Meer kennis op dit grensvlak kan op twee manieren aangewend worden. Het stelt ons in staat om (1) de spel-ervaring van een speler te begrijpen en te beïnvloeden, en (2) de inhoud van een spel aan te passen aan de voorkeuren van de speler. Momenteel bevindt onze kennis van de psychologie die een rol speelt tijdens de interactie tussen speler en computerspel zich in de beginfase. Het is onze bedoeling de psychologische kennis verder te ontwikkelen en nieuwe onderzoeksrichtingen aan te geven.

De probleemstelling van dit proefschrift is: *In welke mate zijn computerspelen een gepast medium voor het meten van individuele verschillen die gebaseerd zijn op psychologische theorieën?* Om deze probleemstelling te onderzoeken bekijken we de incongruentie-theorie en de persoonlijkheidstheorie. We onderzoeken de invloed van beide theorieën op de emotionele uitingen van de spelers en op hun gedrag tijdens het spelen. In totaal voeren we vijf onderzoeken uit om te zien in welke mate de psychologische theorieën het verschil tussen individueën kan verklaren.

In Hoofdstuk 2 presenteren we de beschikbare kennis over de psychologische concepten die relevant zijn voor ons onderzoek, ook presenteren we het onderzoek dat al gedaan is naar ons onderwerp en naar aanverwante onderwerpen.

In Hoofdstuk 3 begint ons onderzoek met onderzoeksvraag 1: *In welke mate zijn computerspelen geschikt voor het meten van incongruentie?* De incongruentie-theorie stelt (kort gezegd) dat spelers zich verveeld zullen voelen bij het spelen van gemakkelijke spelen, dat zij zich plezierig zullen voelen bij het spelen van gebalanceerde spelen, en dat zij zich gefrustreerd zullen voelen bij het spelen van moeilijke spelen. We onderzoeken de relatie tussen (1) de complexiteit van het mentale model van de speler en (2) de emoties die de speler rapporteert. Voor ons onderzoek hebben we een spel gemaakt dat GLOVE heet. In dit spel worden spelers geconfronteerd met een laag-, gebalanceerd- of hoog-complex scenario (m.a.w., een makkelijk, gebalanceerd of moeilijk spel). Uit de resultaten concluderen we dat spelers zich gefrustreerd voelen na moeilijke spellen en zich plezierig voelen na gebalanceerde spellen. We zijn niet in staat geweest te bevestigen dat spelers zich verveeld voelen bij het spelen van gemakkelijke spellen.

Om in ons onderzoek meer focus te leggen op individuele verschillen tussen spelers onderzoeken we in hoofdstuk 4 het psychologische construct extraversie via onderzoeksvraag 2: *In welke mate kunnen complexe psychologische processen zoals extraversie door computerspelen gemeten worden?* We onderzoeken het process extraversie door een scenario te maken voor het spel NEVERWINTER NIGHTS. Voor dit spel herformuleren we verschillende experimenten uit de extraversie-literatuur. In het scenario volgen de spelers een korte verhaallijn terwijl ze verschillende taken uitvoeren. Tijdens het spelen leggen we de reacties van de spelers op de taken vast. Uit de resultaten concluderen we dat, van de 21 vastgelegde taken, er 12 correlleren met extraversie of een facet van extraversie. We denken dat, met een aantal verfijningen, deze aanpak in staat is om grote delen van extraversie en zijn facetten te meten. We meren hierbij op dat extraversie slechts een van de vijf persoonlijkheidstrekken van het "Big-Five" persoonlijkheidsmodel is.

In hoofdstuk 5 breiden we ons onderzoek uit tot de volledige verzameling van vijf persoonlijkheidstrekken en behandelen we onderzoeksvraag 3: *In welke mate kan een data-driven persoonlijkheidsprofiel opgebouwd worden uit spelgedrag?* We onderzoeken alle trekken van het "Big-Five" persoonlijkheidsmodel. We concentreren ons op gedragsmaten die verzameld zijn via een spel in plaats van via reacties op taken in het spel. We kiezen deze aanpak omdat we denken dat deze in meer spellen is toe te passen. We maken een nieuw scenario voor het spel NEVERWINTER NIGHTS. Dit scenario lijkt sterk op scenarios die gebruikt worden in commercieële spellen. We proberen met nadruk we de spelers een grote verscheidenheid aan mogelijke gedragingen te bieden zodat ze vrijelijk hun persoonlijkheid kunnen laten zien. Omdat deze aanpak data-gedreven is creëren we 217 niet-geaggregeerde variabelen die de beweging en het conversatie-gedrag van een speler meten. Ook formuleren we 43 geaggregeerde variabelen waarmee we onze a priori verwachtingen kunnen onderzoeken. We voeren correlatie en lineaire regressie analyses uit voor ieder van de vijf persoonlijkheidstrekken. Uit onze resultaten maken we op dat ieder van de vijf trekken tot uitdrukking komt in het spelgedrag. We zijn dus in staat een data-gedreven persoonlijkheidsprofiel op te bouwen uit spelgedrag. We zien dat een aanpak met zo veel variabelen het risico loopt op "overfitting" met het aantal participanten dat we gebruikt hebben. Om dit risico te verkleinen concentreren we ons in het volgende hoofdstuk op een theorie-gedreven aanpak.

In Hoofdstuk 6 richten we ons op onderzoeksvraag 4: *In welke mate wordt persoonlijkheid in spellen verklaard door een theory-driven model?* Om deze vraag te onderzoeken formuleren we 11 gedragscriteria die gebaseerd zijn op beschrijvingen van gedrag uit literatuur van de "Big-Five" van persoonlijkheidstrekken (openheid, conscientieusheid, extraversie, vriendelijkheid, en neuroticisme). We gebruiken de data die verzameld zijn voor de data-set van Hoofdstuk 5, voor het berekenen van de criteria. Vervolgens doen we een correlatie-analyse tussen de "Big-Five" persoonlijkheidstrekken en de 11 theoretische variabelen. Uit de resultaten kunnen we afleiden dat neuroticisme en vriendelijkheid correleren met drie van de voor hen opgestelde theoretische variabelen. Ook correleert een van de voor conscientieusheid opgestelde variabelen met neuroticisme. Deze aanpak laat zien dat de aannames die gemaakt zijn, met betrekking tot persoonlijkheid en gedrag in computerspelen, goed getest moeten worden voor ze aangenomen worden.

In hoofdstuk 7 beschrijven we validatie onderzoek voor resultaten uit hoofdstuk 5 en 6. Het hoofdstuk behandelt onderzoeksvraag 5: *In welke mate kunnen onze modellen voor* *persoonlijkheid in computerspelen gevalideerd worden in andere spellen?* We onderzoeken of de resultaten van de vorige twee hoofdstukken ook valide zijn in een ander computerspel. Daartoe implementeren we de variabelen uit Hoofdstuk 5 in het startscenario van het commercieel succesvolle spel FALLOUT 3. Na het bijeenbrengen van de data voeren we dezelfde analyses uit die we ook in Hoofdstuk 5 en 6 uitgevoerd hebben. We mogen concluderen dat de resultaten van de data-gedreven aanpak gedeeltelijk gerepliceerd kunnen worden, maar dat dit niet geldt voor de resultaten van de theorie-gedreven aanpak.

In Hoofdstuk 8 presenteren we een discussie over de inzichten die we verkregen hebben tijdens het beantwoorden van de onderzoeksvragen voor deze studie. Het wordt duidelijk dat (1) validatie, (2) voldoende participanten, en (3) de controle variabelen aandachtspunten zijn voor het doen van psychologisch onderzoek in een spel-omgeving. Tevens analyseren we waarom en met welke aanpak we succes hebben gehad bij onze verschillende experimentele manieren van aanpak. We formuleren adviezen voor toekomstige onderzoeken om struikelblokken te vermijden en we vermelden vier gebieden die interessant zijn voor toekomstig onderzoek: (1) vaardigheid, (2) voorkeur, (3) intelligentie, en (4) demografische gegevens.

In Hoofdstuk 9 presenteren we onze conclusies. We geven antwoord op de in Hoofdstuk 1 gestelde onderzoeksvragen en we beantwoorden de probleemstelling. We concluderen dat spellen met succes ingezet kunnen worden om de individuele verschillen tussen spelers te kwantificeren, maar dat we op moeten letten bij (1) analyse, (2) validatie, (3) selectie van het aantal proefpersonen, en (4) de interpretatie van de resultaten.

Curriculum vitae

Giel van Lankveld was born in Gemert, the Netherlands, on April 3, 1982. He studied Informatics on the Hogeschool Zuyd in Heerlen for two years after which he continued to study Psychology at Maastricht University (UM). He received his M.Sc. degree in 2007 with a specialisation in Cognitive Neuroscience. During the period of his psychology study he was active as supervisor for a computer landscape at the university. He was also active in his student society *SV KoKo* as member and later as a president of the Public Relations committee, and member of the almanak committee. Moreover he participated in the startup of the fraternity *Alpha Gamma* (of which he was the first president).

After his graduation, he joined the faculty of Knowledge Engineering at Maastricht University as a Ph.D. candidate, supervised by Prof. dr. H.J. van den Herik, in cooperation (daily supervision) with Dr. ir. P.H.M. Spronck. Later, he joined his supervisors in their move to Tilburg University (TU), to the newly established Tilburg centre for Creative Computing (TiCC). To strengthen the team, Prof. dr. A.R. Arntz was prepared to review the psychological details of his thesis. As a Ph.D. candidate, Giel investigated incongruity and emotion in games, as well as personality in games. He was partly funded by the Korps Landelijke Politie Diensten (KLPD) in the Kennis in Modellen project (KIM). In addition, Giel spent three months as a visiting researcher at the Japan Advanced Institute of Science and Technology (JAIST) under the supervision of Prof. dr. H. Iida. He was also responsible for the maintenance and updates of the website of the International Computer Games Association (ICGA). His research was published in refereed journals, in the proceedings of various (international) conferences. Giel has reviewed papers for the ICGA Journal.

During the study period at Maastricht University Giel was active as a lecturer and practical teacher for the Delphi programming course. During his Ph.D. period at Maastricht University, he was active as a lecturer for the courses Group Dynamics (GD) and Introduction in Modelling (IM). During his Ph.D. period at Tilburg University Giel was active as lecturer for the courses Digital Media and Research Tools (DMRT), Games and AI, and Communication and Information sciences: classics (CIW: classics). Giel has also co-supervised 3 B.Sc. theses and 3 M.Sc. theses, in close cooperation with main supervisor Dr. ir. P.H.M. Spronck.

Publications

1. Bakkes, S. C. J., Spronck, P. H. M., and van Lankveld, G. (2012). Player behavioural modelling for video games. *Entertainment Computing*, 3(3):71–79

2. Joosten, E., van Lankveld, G., and Spronck, P. H. M. (2012). Influencing player emotions using colors. *accepted for publication in the Journal of Intelligent Computing*

3. Spronck, P. H. M., Balemans, I., and van Lankveld, G. (2012). Player profiling with fallout 3. In *Accepted for the AIIDE 2012 conference*

4. Stege, L., van Lankveld, G., and Spronck, P. H. M. (2012). Teaching high school physics with a serious game. *International Journal of Computer Science in Sports*, 11(1)

5. Stege, L., van Lankveld, G., and Spronck, P. H. M. (2011). Teaching electrical engineering with a serious game. In Wiemeyer, J. and Göbel, S., editors, *Theory, Technology & Practice: Proceedings of the GameDays 2011*, pages 29–39

6. van Lankveld, G., Spronck, P. H. M., van den Herik, H. J., and Arntz, A. R. (2011b). Games as personality profiling tools. In *2011 IEEE Conference on Computational Intelligence and Games (CIG)*, pages 197–202. IEEE

7. van Lankveld, G., Schreurs, S., Spronck, P. H. M., and van den Herik, H. J. (2011a). Extraversion in games. In van den Herik, H. J., Plaat, A., and Iida, H., editors, *CG'10 Proceedings of the 7th international conference on Computers and Games*, pages 263–275

8. van Lankveld, G., Spronck, P. H. M., van den Herik, H. J., and Rauterberg, M. (2010). Incongruity-based adaptive game balancing. In *12th International Conference, ACG 2009*, pages 208–221. Springer-Verlag, Heidelberg, Germany

9. Joosten, E., van Lankveld, G., and Spronck, P. H. M. (2010). Colors and emotions in video games. In Ayesh, A., editor, *11th International Conference on Intelligent Games and Simulation GAME-ON 2010*, pages 61–65. EUROSIS

10. van Lankveld, G., Schreurs, S., and Spronck, P. H. M. (2009). Psychologically verified player modelling. In Breitlauch, L., editor, *The 10th International Conference on Intelligent Games and Simulation (GAME-ON 2009)*, pages 12–19

11. van Lankveld, G., Spronck, P. H. M., and Rauterberg, M. (2008). Difficulty scaling through incongruity. In *Proceedings of the Fourth Artificial Intelligence and Interactive Digital Entertainment Conference*, Stanford, California, USA

SIKS dissertation series

1998

1. Johan van den Akker (CWI) *DEGAS - An Active, Temporal Database of Autonomous Objects*

2. Floris Wiesman (UM) *Information Retrieval by Graphically Browsing Meta-Information*

3. Ans Steuten (TUD) *A Contribution to the Linguistic Analysis of Business Conversations within the Language/Action Perspective*

4. Dennis Breuker (UM) *Memory versus Search in Games*

5. E.W.Oskamp (RUL) *Computerondersteuning bij Straftoemeting*

1999

1. Mark Sloof (VU) *Physiology of Quality Change Modelling; Automated modelling of Quality Change of Agricultural Products*

2. Rob Potharst (EUR) *Classification using decision trees and neural nets*

3. Don Beal (UM) *The Nature of Minimax Search*

4. Jacques Penders (UM) *The practical Art of Moving Physical Objects*

5. Aldo de Moor (KUB) *Empowering Communities: A Method for the Legitimate User-Driven Specification of Network Information Systems*

6. Niek J.E. Wijngaards (VU) *Re-design of compositional systems*

7. David Spelt (UT) *Verification support for object database design*

8. Jacques H.J. Lenting (UM) *Informed Gambling: Conception and Analysis of a Multi-Agent Mechanism for Discrete Reallocation*

2000

1. Frank Niessink (VU) *Perspectives on Improving Software Maintenance*

2. Koen Holtman (TUE) *Prototyping of CMS Storage Management*

3. Carolien M.T. Metselaar (UvA) *Sociaal-organisatorische gevolgen van kennistechnologie; een procesbenadering en actorperspectief*

4. Geert de Haan (VU) *ETAG, A Formal Model of Competence Knowledge for User Interface Design*

5. Ruud van der Pol (UM) *Knowledge-based Query Formulation in Information Retrieval*

6. Rogier van Eijk (UU) *Programming Languages for Agent Communication*

7. Niels Peek (UU) *Decision-theoretic Planning of Clinical Patient Management*

8. Veerle Coup (EUR) *Sensitivity Analyis of Decision-Theoretic Networks*

9. Florian Waas (CWI) *Principles of Probabienumerateic Query Optimization*

10. Niels Nes (CWI) *Image Database Management System Design Considerations, Algorithms and Architecture*

11. Jonas Karlsson (CWI) *Scalable Distributed Data Structures for Database Management*

2001

1. Silja Renooij (UU) *Qualitative Approaches to Quantifying Probabienumerateic Networks*

2. Koen Hindriks (UU) *Agent Programming Languages: Programming with Mental Models*

3. Maarten van Someren (UvA) *Learning as problem solving*

4. Evgueni Smirnov (UM) *Conjunctive and Disjunctive Version Spaces with Instance-Based Boundary Sets*

5. Jacco van Ossenbruggen (VU) *Processing Structured Hypermedia: A Matter of Style*

6. Martijn van Welie (VU) *Task-based User Interface Design*

7. Bastiaan Schonhage (VU) *Diva: Architectural Perspectives on Information Visualization*

8. Pascal van Eck (VU) *A Compositional Semantic Structure for Multi-Agent Systems Dynamics*

9. Pieter Jan 't Hoen (RUL) *Towards Distributed Development of Large Object-Oriented Models, Views of Packages as Classes*

10. Maarten Sierhuis (UvA) *Modeling and Simulating Work Practice BRAHMS: a multiagent modeling and simulation language for work practice analysis and design*

11. Tom M. van Engers (VUA) *Knowledge Management: The Role of Mental Models in Business Systems Design*

2002

1. Nico Lassing (VU) *Architecture-Level Modifiability Analysis*

2. Roelof van Zwol (UT) *Modelling and searching web-based document collections*

3. Henk Ernst Blok (UT) *Database Optimization Aspects for Information Retrieval*

4. Juan Roberto Castelo Valdueza (UU) *The Discrete Acyclic Digraph Markov Model in Data Mining*

5. Radu Serban (VU) *The Private Cyberspace Modeling Electronic Environments inhabited by Privacy-concerned Agents*

6. Laurens Mommers (UL) *Applied legal epistemology; Building a knowledge-based ontology of the legal domain*

7. Peter Boncz (CWI) *Monet: A Next-Generation DBMS Kernel For Query-Intensive Applications*

8. Jaap Gordijn (VU) *Value Based Requirements Engineering: Exploring Innovative E-Commerce Ideas*

9. Willem-Jan van den Heuvel(KUB) *Integrating Modern Business Applications with Objectified Legacy Systems*

10. Brian Sheppard (UM) *Towards Perfect Play of Scrabble*

11. Wouter C.A. Wijngaards (VU) *Agent Based Modelling of Dynamics: Biological and Organisational Applications*

12. Albrecht Schmidt (Uva) *Processing XML in Database Systems*

13. Hongjing Wu (TUE) *A Reference Architecture for Adaptive Hypermedia Applications*

14. Wieke de Vries (UU) *Agent Interaction: Abstract Approaches to Modelling, Programming and Verifying Multi-Agent Systems*

15. Rik Eshuis (UT) *Semantics and Verification of UML Activity Diagrams for Workflow Modelling*

16. Pieter van Langen (VU) *The Anatomy of Design: Foundations, Models and Applications*

17. Stefan Manegold (UvA) *Understanding, Modeling, and Improving Main-Memory Database Performance*

2003

1. Heiner Stuckenschmidt (VU) *Ontology-Based Information Sharing in Weakly Structured Environments*

2. Jan Broersen (VU) *Modal Action Logics for Reasoning About Reactive Systems*

3. Martijn Schuemie (TUD) *Human-Computer Interaction and Presence in Virtual Reality Exposure Therapy*

4. Milan Petkovic (UT) *Content-Based Video Retrieval Supported by Database Technology*

5. Jos Lehmann (UvA) *Causation in Artificial Intelligence and Law - A modelling approach*

6. Boris van Schooten (UT) *Development and specification of virtual environments*

7. Machiel Jansen (UvA) *Formal Explorations of Knowledge Intensive Tasks*

8. Yongping Ran (UM) *Repair Based Scheduling*

9. Rens Kortmann (UM) *The resolution of visually guided behaviour*

10. Andreas Lincke (UvT) *Electronic Business Negotiation: Some experimental studies on the interaction between medium, innovation context and culture*

11. Simon Keizer (UT) *Reasoning under Uncertainty in Natural Language Dialogue using Bayesian Networks*

12. Roeland Ordelman (UT) *Dutch speech recognition in multimedia information retrieval*

13. Jeroen Donkers (UM) *Nosce Hostem - Searching with Opponent Models*

14. Stijn Hoppenbrouwers (KUN) *Freezing Language: Conceptualisation Processes across ICT-Supported Organisations*

15. Mathijs de Weerdt (TUD) *Plan Merging in Multi-Agent Systems*

16. Menzo Windhouwer (CWI) *Feature Grammar Systems - Incremental Maintenance of Indexes to Digital Media Warehouses*

17. David Jansen (UT) *Extensions of Statecharts with Probability, Time, and Stochastic Timing*

18. Levente Kocsis (UM) *Learning Search Decisions*

2004

1. Virginia Dignum (UU) *A Model for Organizational Interaction: Based on Agents, Founded in Logic*

2. Lai Xu (UvT) *Monitoring Multi-party Contracts for E-business*

3. Perry Groot (VU) *A Theoretical and Empirical Analysis of Approximation in Symbolic Problem Solving*

4. Chris van Aart (UvA) *Organizational Principles for Multi-Agent Architectures*

5. Viara Popova (EUR) *Knowledge discovery and monotonicity*

6. Bart-Jan Hommes (TUD) *The Evaluation of Business Process Modeling Techniques*

7. Elise Boltjes (UM) *Voorbeeldig onderwijs; voorbeeldgestuurd onderwijs, een opstap naar abstract denken, vooral voor meisjes*

8. Joop Verbeek(UM) *Politie en de Nieuwe Internationale Informatiemarkt, Grensregionale politiële gegevensuitwisseling en digitale expertise*

9. Martin Caminada (VU) *For the Sake of the Argument; explorations into argument-based reasoning*

10. Suzanne Kabel (UvA) *Knowledge-rich indexing of learning-objects*

11. Michel Klein (VU) *Change Management for Distributed Ontologies*

12. The Duy Bui (UT) *Creating emotions and facial expressions for embodied agents*

13. Wojciech Jamroga (UT) *Using Multiple Models of Reality: On Agents who Know how to Play*

14. Paul Harrenstein (UU) *Logic in Conflict. Logical Explorations in Strategic Equilibrium*

15. Arno Knobbe (UU) *Multi-Relational Data Mining*

16. Federico Divina (VU) *Hybrid Genetic Relational Search for Inductive Learning*

17. Mark Winands (UM) *Informed Search in Complex Games*

18. Vania Bessa Machado (UvA) *Supporting the Construction of Qualitative Knowledge Models*

19. Thijs Westerveld (UT) *Using generative probabienumerateic models for multimedia retrieval*

20. Madelon Evers (Nyenrode) *Learning from Design: facilitating multidisciplinary design teams*

2005

1. Floor Verdenius (UvA) *Methodological Aspects of Designing Induction-Based Applications*

2. Erik van der Werf (UM) *AI techniques for the game of Go*

3. Franc Grootjen (RUN) *A Pragmatic Approach to the Conceptualisation of Language*

4. Nirvana Meratnia (UT) *Towards Database Support for Moving Object data*

5. Gabriel Infante-Lopez (UvA) *Two-Level Probabienumerateic Grammars for Natural Language Parsing*

6. Pieter Spronck (UM) *Adaptive Game AI*

7. Flavius Frasincar (TUE) *Hypermedia Presentation Generation for Semantic Web Information Systems*

8. Richard Vdovjak (TUE) *A Model-driven Approach for Building Distributed Ontology-based Web Applications*

9. Jeen Broekstra (VU) *Storage, Querying and Inferencing for Semantic Web Languages*

10. Anders Bouwer (UvA) *Explaining Behaviour: Using Qualitative Simulation in Interactive Learning Environments*

11. Elth Ogston (VU) *Agent Based Matchmaking and Clustering - A Decentralized Approach to Search*

12. Csaba Boer (EUR) *Distributed Simulation in Industry*

13. Fred Hamburg (UL) *Een Computermodel voor het Ondersteunen van Euthanasiebeslissingen*

14. Borys Omelayenko (VU) *Web-Service configuration on the Semantic Web; Exploring how semantics meets pragmatics*

15. Tibor Bosse (VU) *Analysis of the Dynamics of Cognitive Processes*

16. Joris Graaumans (UU) *Usability of XML Query Languages*

17. Boris Shishkov (TUD) *Software Specification Based on Re-usable Business Components*
18. Danielle Sent (UU) *Test-selection strategies for probabienumerateic networks*
19. Michel van Dartel (UM) *Situated Representation*
20. Cristina Coteanu (UL) *Cyber Consumer Law, State of the Art and Perspectives*
21. Wijnand Derks (UT) *Improving Concurrency and Recovery in Database Systems by Exploiting Application Semantics*

2006

1. Samuil Angelov (TUE) *Foundations of B2B Electronic Contracting*
2. Cristina Chisalita (VU) *Contextual issues in the design and use of information technology in organizations*
3. Noor Christoph (UvA) *The role of metacognitive skills in learning to solve problems*
4. Marta Sabou (VU) *Building Web Service Ontologies*
5. Cees Pierik (UU) *Validation Techniques for Object-Oriented Proof Outlines*
6. Ziv Baida (VU) *Software-aided Service Bundling - Intelligent Methods & Tools for Graphical Service Modeling*
7. Marko Smiljanic (UT) *XML schema matching – balancing efficiency and effectiveness by means of clustering*
8. Eelco Herder (UT) *Forward, Back and Home Again - Analyzing User behaviour on the Web*
9. Mohamed Wahdan (UM) *Automatic Formulation of the Auditor's Opinion*
10. Ronny Siebes (VU) *Semantic Routing in Peer-to-Peer Systems*
11. Joeri van Ruth (UT) *Flattening Queries over Nested Data Types*
12. Bert Bongers (VU) *Interactivation - Towards an e-cology of people, our technological environment, and the arts*
13. Henk-Jan Lebbink (UU) *Dialogue and Decision Games for Information Exchanging Agents*
14. Johan Hoorn (VU) *Software Requirements: Update, Upgrade, Redesign - towards a Theory of Requirements Change*
15. Rainer Malik (UU) *CONAN: Text Mining in the Biomedical Domain*
16. Carsten Riggelsen (UU) *Approximation Methods for Efficient Learning of Bayesian Networks*
17. Stacey Nagata (UU) *User Assistance for Multitasking with Interruptions on a Mobile Device*
18. Valentin Zhizhkun (UvA) *Graph transformation for Natural Language Processing*
19. Birna van Riemsdijk (UU) *Cognitive Agent Programming: A Semantic Approach*
20. Marina Velikova (UvT) *Monotone models for prediction in data mining*
21. Bas van Gils (RUN) *Aptness on the Web*
22. Paul de Vrieze (RUN) *Fundaments of Adaptive Personalisation*
23. Ion Juvina (UU) *Development of Cognitive Model for Navigating on the Web*
24. Laura Hollink (VU) *Semantic Annotation for Retrieval of Visual Resources*
25. Madalina Drugan (UU) *Conditional log-likelihood MDL and Evolutionary MCMC*
26. Vojkan Mihajlovic (UT) *Score Region Algebra: A Flexible Framework for Structured Information Retrieval*
27. Stefano Bocconi (CWI) *Vox Populi: generating video documentaries from semantically annotated media repositories*
28. Borkur Sigurbjornsson (UvA) *Focused Information Access using XML Element Retrieval*

2007

1. Kees Leune (UvT) *Access Control and Service-Oriented Architectures*
2. Wouter Teepe (RUG) *Reconciling Information Exchange and Confidentiality: A Formal Approach*
3. Peter Mika (VU) *Social Networks and the Semantic Web*
4. Jurriaan van Diggelen (UU) *Achieving Semantic Interoperability in Multi-agent Systems: a dialogue-based approach*
5. Bart Schermer (UL) *Software Agents, Surveillance, and the Right to Privacy: a Legislative Framework for Agent-enabled Surveillance*
6. Gilad Mishne (UvA) *Applied Text Analytics for Blogs*
7. Natasa Jovanovic' (UT) *To Whom It May Concern - Addressee Identification in Face-to-Face Meetings*
8. Mark Hoogendoorn (VU) *Modeling of Change in Multi-Agent Organizations*

9. David Mobach (VU) *Agent-Based Mediated Service Negotiation*

10. Huib Aldewereld (UU) *Autonomy vs. Conformity: an Institutional Perspective on Norms and Protocols*

11. Natalia Stash (TUE) *Incorporating Cognitive/Learning Styles in a General-Purpose Adaptive Hypermedia System*

12. Marcel van Gerven (RUN) *Bayesian Networks for Clinical Decision Support: A Rational Approach to Dynamic Decision-Making under Uncertainty*

13. Rutger Rienks (UT) *Meetings in Smart Environments; Implications of Progressing Technology*

14. Niek Bergboer (UM) *Context-Based Image Analysis*

15. Joyca Lacroix (UM) *NIM: a Situated Computational Memory Model*

16. Davide Grossi (UU) *Designing Invisible Handcuffs. Formal investigations in Institutions and Organizations for Multi-agent Systems*

17. Theodore Charitos (UU) *Reasoning with Dynamic Networks in Practice*

18. Bart Orriens (UvT) *On the development an management of adaptive business collaborations*

19. David Levy (UM) *Intimate relationships with artificial partners*

20. Slinger Jansen (UU) *Customer Configuration Updating in a Software Supply Network*

21. Karianne Vermaas (UU) *Fast diffusion and broadening use: A research on residential adoption and usage of broadband internet in the Netherlands between 2001 and 2005*

22. Zlatko Zlatev (UT) *Goal-oriented design of value and process models from patterns*

23. Peter Barna (TUE) *Specification of Application Logic in Web Information Systems*

24. Georgina Ramrez Camps (CWI) *Structural Features in XML Retrieval*

25. Joost Schalken (VU) *Empirical Investigations in Software Process Improvement*

2008

1. Katalin Boer-Sorbn (EUR) *Agent-Based Simulation of Financial Markets: A modular,continuous-time approach*

2. Alexei Sharpanskykh (VU) *On Computer-Aided Methods for Modeling and Analysis of Organizations*

3. Vera Hollink (UvA) *Optimizing hierarchical menus: a usage-based approach*

4. Ander de Keijzer (UT) *Management of Uncertain Data - towards unattended integration*

5. Bela Mutschler (UT) *Modeling and simulating causal dependencies on process-aware information systems from a cost perspective*

6. Arjen Hommersom (RUN) *On the Application of Formal Methods to Clinical Guidelines, an Artificial Intelligence Perspective*

7. Peter van Rosmalen (OU) *Supporting the tutor in the design and support of adaptive e-learning*

8. Janneke Bolt (UU) *Bayesian Networks: Aspects of Approximate Inference*

9. Christof van Nimwegen (UU) *The paradox of the guided user: assistance can be counter-effective*

10. Wauter Bosma (UT) *Discourse oriented summarization*

11. Vera Kartseva (VU) *Designing Controls for Network Organizations: A Value-Based Approach*

12. Jozsef Farkas (RUN) *A Semiotically Oriented Cognitive Model of Knowledge Representation*

13. Caterina Carraciolo (UvA) *Topic Driven Access to Scientific Handbooks*

14. Arthur van Bunningen (UT) *Context-Aware Querying; Better Answers with Less Effort*

15. Martijn van Otterlo (UT) *The Logic of Adaptive behaviour: Knowledge Representation and Algorithms for the Markov Decision Process Framework in First-Order Domains*

16. Henriette van Vugt (VU) *Embodied agents from a user's perspective*

17. Martin Op 't Land (TUD) *Applying Architecture and Ontology to the Splitting and Allying of Enterprises*

18. Guido de Croon (UM) *Adaptive Active Vision*

19. Henning Rode (UT) *From Document to Entity Retrieval: Improving Precision and Performance of Focused Text Search*

20. Rex Arendsen (UvA) *Geen bericht, goed bericht. Een onderzoek naar de effecten van de introductie van elektronisch berichtenverkeer met de overheid op de administratieve lasten van bedrijven*
21. Krisztian Balog (UvA) *People Search in the Enterprise*
22. Henk Koning (UU) *Communication of IT-Architecture*
23. Stefan Visscher (UU) *Bayesian network models for the management of ventilator-associated pneumonia*
24. Zharko Aleksovski (VU) *Using background knowledge in ontology matching*
25. Geert Jonker (UU) *Efficient and Equitable Exchange in Air Traffic Management Plan Repair using Spender-signed Currency*
26. Marijn Huijbregts (UT) *Segmentation, Diarization and Speech Transcription: Surprise Data Unraveled*
27. Hubert Vogten (OU) *Design and Implementation Strategies for IMS Learning Design*
28. Ildiko Flesch (RUN) *On the Use of Independence Relations in Bayesian Networks*
29. Dennis Reidsma (UT) *Annotations and Subjective Machines - Of Annotators, Embodied Agents, Users, and Other Humans*
30. Wouter van Atteveldt (VU) *Semantic Network Analysis: Techniques for Extracting, Representing and Querying Media Content*
31. Loes Braun (UM) *Pro-Active Medical Information Retrieval*
32. Trung H. Bui (UT) *Toward Affective Dialogue Management using Partially Observable Markov Decision Processes*
33. Frank Terpstra (UvA) *Scientific Workflow Design; theoretical and practical issues*
34. Jeroen de Knijf (UU) *Studies in Frequent Tree Mining*
35. Ben Torben Nielsen (UvT) *Dendritic morphologies: function shapes structure*

2009

1. Rasa Jurgelenaite (RUN) *Symmetric Causal Independence Models*
2. Willem Robert van Hage (VU) *Evaluating Ontology-Alignment Techniques*
3. Hans Stol (UvT) *A Framework for Evidence-based Policy Making Using IT*
4. Josephine Nabukenya (RUN) *Improving the Quality of Organisational Policy Making using Collaboration Engineering*
5. Sietse Overbeek (RUN) *Bridging Supply and Demand for Knowledge Intensive Tasks - Based on Knowledge, Cognition, and Quality*
6. Muhammad Subianto (UU) *Understanding Classification*
7. Ronald Poppe (UT) *Discriminative Vision-Based Recovery and Recognition of Human Motion*
8. Volker Nannen (VU) *Evolutionary Agent-Based Policy Analysis in Dynamic Environments*
9. Benjamin Kanagwa (RUN) *Design, Discovery and Construction of Service-oriented Systems*
10. Jan Wielemaker (UvA) *Logic programming for knowledge-intensive interactive applications*
11. Alexander Boer (UvA) *Legal Theory, Sources of Law & the Semantic Web*
12. Peter Massuthe (TUE, Humboldt-Universitaet zu Berlin) *Operating Guidelines for Services*
13. Steven de Jong (UM) *Fairness in Multi-Agent Systems*
14. Maksym Korotkiy (VU) *From ontology-enabled services to service-enabled ontologies (making ontologies work in e-science with ONTO-SOA)*
15. Rinke Hoekstra (UvA) *Ontology Representation - Design Patterns and Ontologies that Make Sense*
16. Fritz Reul (UvT) *New Architectures in Computer Chess*
17. Laurens van der Maaten (UvT) *Feature Extraction from Visual Data*
18. Fabian Groffen (CWI) *Armada, An Evolving Database System*
19. Valentin Robu (CWI) *Modeling Preferences, Strategic Reasoning and Collaboration in Agent-Mediated Electronic Markets*
20. Bob van der Vecht (UU) *Adjustable Autonomy: Controling Influences on Decision Making*
21. Stijn Vanderlooy (UM) *Ranking and Reliable Classification*
22. Pavel Serdyukov (UT) *Search For Expertise: Going beyond direct evidence*
23. Peter Hofgesang (VU) *Modelling Web Usage in a Changing Environment*
24. Annerieke Heuvelink (VUA) *Cognitive Models for Training Simulations*

25. Alex van Ballegooij (CWI) *"RAM: Array Database Management through Relational Mapping"*

26. Fernando Koch (UU) *An Agent-Based Model for the Development of Intelligent Mobile Services*

27. Christian Glahn (OU) *Contextual Support of social Engagement and Reflection on the Web*

28. Sander Evers (UT) *Sensor Data Management with Probabienumerateic Models*

29. Stanislav Pokraev (UT) *Model-Driven Semantic Integration of Service-Oriented Applications*

30. Marcin Zukowski (CWI) *Balancing vectorized query execution with bandwidth-optimized storage*

31. Sofiya Katrenko (UvA) *A Closer Look at Learning Relations from Text*

32. Rik Farenhorst (VU) and Remco de Boer (VU) *Architectural Knowledge Management: Supporting Architects and Auditors*

33. Khiet Truong (UT) *How Does Real Affect Affect Affect Recognition In Speech?*

34. Inge van de Weerd (UU) *Advancing in Software Product Management: An Incremental Method Engineering Approach*

35. Wouter Koelewijn (UL) *Privacy en Politiegegevens; Over geautomatiseerde normatieve informatie-uitwisseling*

36. Marco Kalz (OUN) *Placement Support for Learners in Learning Networks*

37. Hendrik Drachsler (OUN) *Navigation Support for Learners in Informal Learning Networks*

38. Riina Vuorikari (OU) *Tags and self-organisation: a metadata ecology for learning resources in a multilingual context*

39. Christian Stahl (TUE, Humboldt-Universitaet zu Berlin) *Service Substitution – A behavioural Approach Based on Petri Nets*

40. Stephan Raaijmakers (UvT) *Multinomial Language Learning: Investigations into the Geometry of Language*

41. Igor Berezhnyy (UvT) *Digital Analysis of Paintings*

42. Toine Bogers *Recommender Systems for Social Bookmarking*

43. Virginia Nunes Leal Franqueira (UT) *Finding Multi-step Attacks in Computer Networks using Heuristic Search and Mobile Ambients*

44. Roberto Santana Tapia (UT) *Assessing Business-IT Alignment in Networked Organizations*

45. Jilles Vreeken (UU) *Making Pattern Mining Useful*

46. Loredana Afanasiev (UvA) *Querying XML: Benchmarks and Recursion*

2010

1. Matthijs van Leeuwen (UU) *Patterns that Matter*

2. Ingo Wassink (UT) *Work flows in Life Science*

3. Joost Geurts (CWI) *A Document Engineering Model and Processing Framework for Multimedia documents*

4. Olga Kulyk (UT) *Do You Know What I Know? Situational Awareness of Co-located Teams in Multidisplay Environments*

5. Claudia Hauff (UT) *Predicting the Effectiveness of Queries and Retrieval Systems*

6. Sander Bakkes (UvT) *Rapid Adaptation of Video Game AI*

7. Wim Fikkert (UT) *Gesture interaction at a Distance*

8. Krzysztof Siewicz (UL) *Towards an Improved Regulatory Framework of Free Software. Protecting user freedoms in a world of software communities and eGovernments*

9. Hugo Kielman (UL) *A Politiele gegevensverwerking en Privacy, Naar een effectieve waarborging*

10. Rebecca Ong (UL) *Mobile Communication and Protection of Children*

11. Adriaan Ter Mors (TUD) *The world according to MARP: Multi-Agent Route Planning*

12. Susan van den Braak (UU) *Sensemaking software for crime analysis*

13. Gianluigi Folino (RUN) *High Performance Data Mining using Bio-inspired techniques*

14. Sander van Splunter (VU) *Automated Web Service Reconfiguration*

15. Lianne Bodenstaff (UT) *Managing Dependency Relations in Inter-Organizational Models*

16. Sicco Verwer (TUD) *Efficient Identification of Timed Automata, theory and practice*

17. Spyros Kotoulas (VU) *Scalable Discovery of Networked Resources: Algorithms, Infrastructure, Applications*

18. Charlotte Gerritsen (VU) *Caught in the Act: Investigating Crime by Agent-Based Simulation*

19. Henriette Cramer (UvA) *People's Responses to Autonomous and Adaptive Systems*

20. Ivo Swartjes (UT) *Whose Story Is It Anyway? How Improv Informs Agency and Authorship of Emergent Narrative*

21. Harold van Heerde (UT) *Privacy-aware data management by means of data degradation*

22. Michiel Hildebrand (CWI) *End-user Support for Access to Heterogeneous Linked Data*

23. Bas Steunebrink (UU) *The Logical Structure of Emotions*

24. Dmytro Tykhonov *Designing Generic and Efficient Negotiation Strategies*

25. Zulfiqar Ali Memon (VU) *Modelling Human-Awareness for Ambient Agents: A Human Mindreading Perspective*

26. Ying Zhang (CWI) *XRPC: Efficient Distributed Query Processing on Heterogeneous XQuery Engines*

27. Marten Voulon (UL) *Automatisch contracteren*

28. Arne Koopman (UU) *Characteristic Relational Patterns*

29. Stratos Idreos(CWI) *Database Cracking: Towards Auto-tuning Database Kernels*

30. Marieke van Erp (UvT) *Accessing Natural History - Discoveries in data cleaning, structuring, and retrieval*

31. Victor de Boer (UvA) *Ontology Enrichment from Heterogeneous Sources on the Web*

32. Marcel Hiel (UvT) *An Adaptive Service Oriented Architecture: Automatically solving Interoperability Problems*

33. Robin Aly (UT) *Modeling Representation Uncertainty in Concept-Based Multimedia Retrieval*

34. Teduh Dirgahayu (UT) *Interaction Design in Service Compositions*

35. Dolf Trieschnigg (UT) *Proof of Concept: Concept-based Biomedical Information Retrieval*

36. Jose Janssen (OU) *Paving the Way for Lifelong Learning; Facilitating competence development through a learning path specification*

37. Niels Lohmann (TUE) *Correctness of services and their composition*

38. Dirk Fahland (TUE) *From Scenarios to components*

39. Ghazanfar Farooq Siddiqui (VU) *Integrative modeling of emotions in virtual agents*

40. Mark van Assem (VU) *Converting and Integrating Vocabularies for the Semantic Web*

41. Guillaume Chaslot (UM) *Monte-Carlo Tree Search*

42. Sybren de Kinderen (VU) *Needs-driven service bundling in a multi-supplier setting - the computational e3-service approach*

43. Peter van Kranenburg (UU) *A Computational Approach to Content-Based Retrieval of Folk Song Melodies*

44. Pieter Bellekens (TUE) *An Approach towards Context-sensitive and User-adapted Access to Heterogeneous Data Sources, Illustrated in the Television Domain*

45. Vasilios Andrikopoulos (UvT) *A theory and model for the evolution of software services*

46. Vincent Pijpers (VU) *e3alignment: Exploring Inter-Organizational Business-ICT Alignment*

47. Chen Li (UT) *Mining Process Model Variants: Challenges, Techniques, Examples*

48. Milan Lovric (EUR) *behavioural Finance and Agent-Based Artificial Markets*

49. Jahn-Takeshi Saito (UM) *Solving difficult game positions*

50. Bouke Huurnink (UvA) *Search in Audiovisual Broadcast Archives*

51. Alia Khairia Amin (CWI) *Understanding and supporting information seeking tasks in multiple sources*

52. Peter-Paul van Maanen (VU) *Adaptive Support for Human-Computer Teams: Exploring the Use of Cognitive Models of Trust and Attention*

53. Edgar Meij (UvA) *Combining Concepts and Language Models for Information Access*

2011

1. Botond Cseke (RUN) *Variational Algorithms for Bayesian Inference in Latent Gaussian Models*
2. Nick Tinnemeier(UU) *Work flows in Life Science*
3. Jan Martijn van der Werf (TUE) *Compositional Design and Verification of Component-Based Information Systems*
4. Hado van Hasselt (UU) *Insights in Reinforcement Learning; Formal analysis and empirical evaluation of temporal-difference learning algorithms*
5. Base van der Raadt (VU) *Enterprise Architecture Coming of Age - Increasing the Performance of an Emerging Discipline*
6. Yiwen Wang (TUE) *Semantically-Enhanced Recommendations in Cultural Heritage*
7. Yujia Cao (UT) *Multimodal Information Presentation for High Load Human Computer Interaction*
8. Nieske Vergunst (UU) *BDI-based Generation of Robust Task-Oriented Dialogues*
9. Tim de Jong (OU) *Contextualised Mobile Media for Learning*
10. Bart Bogaert (UvT) *Cloud Content Contention*
11. Dhaval Vyas (UT) *Designing for Awareness: An Experience-focused HCI Perspective*
12. Carmen Bratosin (TUE) *Grid Architecture for Distributed Process Mining*
13. Xiaoyu Mao (UvT) *Airport under Control. Multiagent Scheduling for Airport Ground Handling*
14. Milan Lovric (EUR) *behavioural Finance and Agent-Based Artificial Markets*
15. Marijn Koolen (UvA) *The Meaning of Structure: the Value of Link Evidence for Information Retrieval*
16. Maarten Schadd (UM) *Selective Search in Games of Different Complexity*
17. Jiyin He (UvA) *Exploring Topic Structure: Coherence, Diversity and Relatedness*
18. Mark Ponsen (UM) *Strategic Decision-Making in complex games*
19. Ellen Rusman (OU) *The Mind ' s Eye on Personal Profiles*
20. Qing Gu (VU) *Guiding service-oriented software engineering - A view-based approach*
21. Linda Terlouw (TUD) *Modularization and Specification of Service-Oriented Systems*
22. Junte Zhang (UvA) *System Evaluation of Archival Description and Access*
23. Wouter Weerkamp (UvA) *Finding People and their Utterances in Social Media*
24. Herwin van Welbergen (UT) *Behavior Generation for Interpersonal Coordination with Virtual Humans On Specifying, Scheduling and Realizing Multimodal Virtual Human Behavior*
25. Syed Waqar ul Qounain Jaffry (VU) *Analysis and Validation of Models for Trust Dynamics*
26. Matthijs Aart Pontier (VU) *Virtual Agents for Human Communication - Emotion Regulation and Involvement-Distance Trade-Offs in Embodied Conversational Agents and Robots*
27. Aniel Bhulai (VU) *Dynamic website optimization through autonomous management of design patterns*
28. Rianne Kaptein(UvA) *Effective Focused Retrieval by Exploiting Query Context and Document Structure*
29. Faisal Kamiran (TUE) *Discrimination-aware Classification*
30. Egon van den Broek (UT) *Affective Signal Processing (ASP): Unraveling the mystery of emotions*
31. Ludo Waltman (EUR) *Computational and Game-Theoretic Approaches for Modeling Bounded Rationality*
32. Nees-Jan van Eck (EUR) *Methodological Advances in Bibliometric Mapping of Science*
33. Tom van der Weide (UU) *Arguing to Motivate Decisions*
34. Paolo Turrini (UU) *Strategic Reasoning in Interdependence: Logical and Game-theoretical Investigations*
35. Maaike Harbers (UU) *Explaining Agent Behavior in Virtual Training*
36. Erik van der Spek (UU) *Experiments in serious game design: a cognitive approach*
37. Adriana Burlutiu (RUN) *Machine Learning for Pairwise Data, Applications for Preference Learning and Supervised Network Inference*
38. Nyree Lemmens (UM) *Bee-inspired Distributed Optimization*
39. Joost Westra (UU) *Organizing Adaptation using Agents in Serious Games*
40. Viktor Clerc (VU) *Architectural Knowledge Management in Global Software Development*

41. Luan Ibraimi (UT) *Cryptographically Enforced Distributed Data Access Control*

42. Michal Sindlar (UU) *Explaining Behavior through Mental State Attribution*

43. Henk van der Schuur (UU) *Process Improvement through Software Operation Knowledge*

44. Boris Reuderink (UT) *Robust Brain-Computer Interfaces*

45. Herman Stehouwer (UvT) *Statistical Language Models for Alternative Sequence Selection*

46. Beibei Hu (TUD) *Towards Contextualized Information Delivery: A Rule-based Architecture for the Domain of Mobile Police Work*

47. Azizi Bin Ab Aziz(VU) *Exploring Computational Models for Intelligent Support of Persons with Depression*

48. Mark Ter Maat (UT) *Response Selection and Turn-taking for a Sensitive Artificial Listening Agent*

49. Andreea Niculescu (UT) *Conversational interfaces for task-oriented spoken dialogues: design aspects influencing interaction quality*

2012

1. Terry Kakeeto (UvT) *Relationship Marketing for SMEs in Uganda*

2. Muhammad Umair(VU) *Adaptivity, emotion, and Rationality in Human and Ambient Agent Models*

3. Adam Vanya (VU) *Supporting Architecture Evolution by Mining Software Repositories*

4. Jurriaan Souer (UU) *Development of Content Management System-based Web Applications*

5. Marijn Plomp (UU) *Maturing Interorganisational Information Systems*

6. Wolfgang Reinhardt (OU) *Awareness Support for Knowledge Workers in Research Networks*

7. Rianne van Lambalgen (VU) *When the Going Gets Tough: Exploring Agent-based Models of Human Performance under Demanding Conditions*

8. Gerben de Vries (UvA) *Kernel Methods for Vessel Trajectories*

9. Ricardo Neisse (UT) *Trust and Privacy Management Support for Context-Aware Service Platforms*

10. David Smits (TUE) *Towards a Generic Distributed Adaptive Hypermedia Environment*

11. J.C.B. Rantham Prabhakara (TUE) *Process Mining in the Large: Preprocessing, Discovery, and Diagnostics*

12. Kees van der Sluijs (TUE) *Model Driven Design and Data Integration in Semantic Web Information Systems*

13. Suleman Shahid (UvT) *Fun and Face: Exploring non-verbal expressions of emotion during playful interactions*

14. Evgeny Knutov(TUE) *Generic Adaptation Framework for Unifying Adaptive Web-based Systems*

15. Natalie van der Wal (VU) *Social Agents. Agent-Based Modelling of Integrated Internal and Social Dynamics of Cognitive and Affective Processes*

16. Fiemke Both (VU) *Helping people by understanding them - Ambient Agents supporting task execution and depression treatment*

17. Amal Elgammal (UvT) *Towards a Comprehensive Framework for Business Process Compliance*

18. Eltjo Poort (VU) *Improving Solution Architecting Practices*

19. Helen Schonenberg (TUE) *What's Next? Operational Support for Business Process Execution*

20. Ali Bahramisharif (RUN) *Covert Visual Spatial Attention, a Robust Paradigm for Brain-Computer Interfacing*

21. Roberto Cornacchia (TUD) *Querying Sparse Matrices for Information Retrieval*

22. Thijs Vis (UvT) *Intelligence, politie en veiligheidsdienst: verenigbare grootheden?*

23. Christian Muehl (UT) *Toward Affective Brain-Computer Interfaces: Exploring the Neurophysiology of Affect during Human Media Interaction*

24. Laurens van der Werff (UT) *Evaluation of Noisy Transcripts for Spoken Document Retrieval*

25. Silja Eckartz (UT) *Managing the Business Case Development in Inter-Organizational IT Projects: A Methodology and its Application*

26. Emile de Maat (UvA) *Making Sense of Legal Text*

27. Hayrettin Gurkok (UT) *Mind the Sheep! User Experience Evaluation & Brain-Computer Interface Games*

28. Nancy Pascall (UvT) *Engendering Technology Empowering Women*

29. Almer Tigelaar (UT) *Peer-to-Peer Information Retrieval*

30. Alina Pommeranz (TUD) *Designing Human-Centered Systems for Reflective Decision Making*

31. Emily Bagarukayo (RUN) *A Learning by Construction Approach for Higher Order Cognitive Skills Improvement, Building Capacity and Infrastructure*

32. Wietske Visser (TUD) *Qualitative multi-criteria preference representation and reasoning*

33. Rory Sie (OUN) *Coalitions in Cooperation Networks (COCOON)*

34. Pavol Jancura (RUN) *Evolutionary analysis in PPI networks and applications*

35. Evert Haasdijk (VU) *Never Too Old To Learn – On-line Evolution of Controllers in Swarm- and Modular Robotics*

36. Denis Ssebugwawo (RUN) *Analysis and Evaluation of Collaborative Modeling Processes*

37. Agnes Nakakawa (RUN) *A Collaboration Process for Enterprise Architecture Creation*

TiCC Ph.D. series

1. Pashiera Barkhuysen. *Audiovisual Prosody in Interaction.* Promotores: M.G.J. Swerts, E.J. Krahmer. Tilburg, 3 October 2008.

2. Ben Torben-Nielsen. *Dendritic morphology: function shapes structure.* Promotores: H.J. van den Herik, E.O. Postma. Co-promotor: K.P. Tuyls. Tilburg, 3 December 2008.

3. Hans Stol. *A framework for evidence-based policy making using IT.* Promotor: H.J. van den Herik. Tilburg, 21 January 2009.

4. Jeroen Geertzen. *Dialogue act recognition and prediction.* Promotor: H. Bunt. Co-promotor: J.M.B. Terken. Tilburg, 11 February 2009.

5. Sander Canisius. *Structured prediction for natural language processing.* Promotores: A.P.J. van den Bosch, W. Daelemans. Tilburg, 13 February 2009.

6. Fritz Reul. *New Architectures in Computer Chess.* Promotor: H.J. van den Herik. Co-promotor: J.W.H.M. Uiterwijk. Tilburg, 17 June 2009.

7. Laurens van der Maaten. *Feature Extraction from Visual Data.* Promotores: E.O. Postma, H.J. van den Herik. Co-promotor: A.G. Lange. Tilburg, 23 June 2009 (cum laude).

8. Stephan Raaijmakers. *Multinomial Language Learning.* Promotores: W. Daelemans, A.P.J. van den Bosch. Tilburg, 1 December 2009.

9. Igor Berezhnoy. *Digital Analysis of Paintings.* Promotores: E.O. Postma, H.J. van den Herik. Tilburg, 7 December 2009.

10. Toine Bogers. *Recommender Systems for Social Bookmarking.* Promotor: A.P.J. van den Bosch. Tilburg, 8 December 2009.

11. Sander Bakkes. *Rapid Adaptation of Video Game AI.* Promotor: H.J. van den Herik. Co-promotor: P. Spronck. Tilburg, 3 March 2010.

12. Maria Mos. *Complex Lexical Items.* Promotor: A.P.J. van den Bosch. Co-promotores: A. Vermeer, A. Backus. Tilburg, 12 May 2010 (in collaboration with the Department of Language and Culture Studies).

13. Marieke van Erp. *Accessing Natural History. Discoveries in data cleaning, structuring, and retrieval.* Promotor: A.P.J. van den Bosch. Co-promotor: P.K. Lendvai. Tilburg, 30 June 2010.

14. Edwin Commandeur. *Implicit Causality and Implicit Consequentiality in Language Comprehension.* Promotores: L.G.M. Noordman, W. Vonk. Co-promotor: R. Cozijn. Tilburg, 30 June 2010.

15. Bart Bogaert. *Cloud Content Contention.* Promotores: H.J. van den Herik, E.O. Postma. Tilburg, 30 March 2011.

16. Xiaoyu Mao. *Airport under Control.* Promotores: H.J. van den Herik, E.O. Postma. Co-promotores: N. Roos, A. Salden. Tilburg, 25 May 2011.

17. Olga Petukhova. *Multidimensional Dialogue Modelling.* Promotor: H. Bunt. Tilburg, 1 September 2011.

18. Lisette Mol. *Language in the Hands.* Promotores: E.J. Krahmer, F. Maes, M.G.J. Swerts. Tilburg, 7 November 2011 (cum laude).

19. Herman Stehouwer. *Statistical Language Models for Alternative Sequence Selection.* Promotores: A.P.J. van den Bosch, H.J. van den Herik. Co-Promotor: M.M. van Zaanen. Tilburg, 7 December 2011.

20. Terry Kakeeto-Aelen. *Relationship Marketing for SMEs in Uganda.* Promotores: J.Chr. van Dalen, H.J. van den Herik. Co-promotor: B.A. Van de Walle. Tilburg, 2012.

21. Suleman Shahid. *Fun & Face: Exploring non-verbal expressions of emotion during playful interactions.* Promotores: E.J. Krahmer, M.G.J. Swerts. Tilburg, 25 May 2012.

22. Thijs Vis. *Intelligence, Politie en Veiligheidsdienst: Verenigbare Grootheden?* Promotores: T.A. de Roos, H.J. van den Herik, A.C.M. Spapens. Tilburg, 6 June 2012.

23. Nancy Pascall. *Engendering Technology Empowering Women.* Promotores: H.J. van den Herik, M. Diocaretz. Tilburg, 19 November 2012.

24. Agus Gunawan. *Information Access for SMEs in Indonesia.* Promotor: H.J. van den Herik. Co-promotores: M. Wahdan, B.A. Van de Walle. Tilburg, 19 December 2012.

25. Giel van Lankveld. *Quantifying Individual Player Differences.* Promotores: H.J. van den Herik, A.R. Arntz. Co-promotor: P. Spronck. Tilburg, 27 Febuary 2013.