

Title	The placebo effect: measurement by multiple methods
Authors	Morrisroe, Richard Patrick
Publication date	2015
Original Citation	Morrisroe, R. P. 2015. The placebo effect: measurement by multiple methods. PhD Thesis, University College Cork.
Type of publication	Doctoral thesis
Rights	© 2015, Richard P. Morrisroe. - http://creativecommons.org/licenses/by-nc-nd/3.0/
Download date	2026-08-11 11:24:42
Item downloaded from	https://hdl.handle.net/10468/2106

The Placebo Effect: Measurement by Multiple Methods

Richard Patrick Morrisroe
BA (HONS)
99473437



NATIONAL UNIVERSITY OF IRELAND, CORK

SCHOOL OF APPLIED PSYCHOLOGY

**Thesis submitted for the degree of
Doctor of Philosophy**

3 June 2015

Head of School: Prof John McCarthy

Supervisors: Dr Z. Di Blasi
Dr Sean Hammond

Research supported by Irish Research Council for Humanities and Social Sciences

Contents

List of Figures	vii
List of Tables	ix
Abstract	xv
Acknowledgements	xvi
1 Introduction	1
1.1 Preface	1
1.2 Introduction to Placebo & Implicit Measures	3
1.2.1 Placebo	3
1.2.2 Treatment Expectancies	4
1.2.3 Moving Beyond Self-Report	6
1.2.4 Older non-self report methods	7
1.2.5 Modern Indirect Measures	7
1.2.6 Introduction to the Implicit Association Test (IAT)	8
1.2.7 Description of the Procedure	8
1.3 Structure of this Thesis	10
2 Literature Review	11
2.1 Introduction to this review	12
2.2 The Concept of the Placebo	12
2.2.1 History of the Concept	12
2.2.2 Placebos in Randomised Trials	14
2.2.3 Problems with the clinical definitions of placebo	15
2.2.4 Definition for this thesis	18
2.3 Theories of Placebo Effects	18
2.3.1 Conditioning	18
2.3.2 Expectancies	20
2.3.2.1 Expectancy Manipulations	20
2.3.3 Motivational Theories	22
2.4 Moderators of the Placebo Effect	24
2.4.1 Certain and Uncertain Expectations	24
2.4.2 Treatment Preference	25
2.4.3 Price of placebo	25
2.4.4 Patient/Participant Characteristics	26
2.4.4.1 Somatic Focus	26
2.4.5 Provider Factors	27
2.5 Effect Sizes for Placebo Studies	28
2.6 Physiology of the Placebo Effect	29
2.6.1 Opioids and Placebo	29
2.6.1.1 Placebo and Skin Conductance	30
2.7 Review of the Placebo Effect	31
2.8 The Implicit Association Test	33
2.9 Psychometric Analysis of the IAT	33
2.9.1 Convergent Validity	33
2.9.1.1 Relationships with Explicit Measures	33
2.9.2 Predictive Validity of the IAT	35
2.9.3 Reliability of the IAT	37
2.10 Moderators of the IAT	38

2.10.1	Contextual Moderators of IAT Effects	38
2.10.2	Cognitive Moderators of IAT effects	38
2.11	Conclusions	39
2.12	Placebo: Implicit and Explicit Expectancies	40
2.12.1	The research evidence for the combination of the two measures	41
3	Theory & Methodology	44
3.1	Introduction	45
3.2	Introduction to the theory	45
3.2.1	Implicit Measures and Dual Process Models of Mind	46
3.2.2	Conscious Conceptions of the Placebo	46
3.2.3	Embodied Cognition and Placebo	47
3.2.4	Additive Models of Placebo	48
3.2.5	Social Aspects of Placebo	49
3.2.6	Towards an Embodied Conception of Placebo	50
3.3	Development of the IAT	52
3.3.1	Repertory Grids	52
3.3.2	Development of IAT	53
3.4	Quantitative Research Methodology	53
3.4.1	Problems of Sample Inference	54
3.4.1.1	The Problem	54
3.4.1.2	A Solution	55
3.4.1.3	Cross-Validation	55
3.4.1.4	Regression Models	56
3.4.2	Psychometric Analyses	58
3.4.2.1	Factor Analysis	58
3.4.2.2	Item Response Theory	60
3.4.2.3	Structural equation modelling	62
3.4.3	General Analytical Approach	63
3.5	Conclusions	65
4	Health Optimism and Mindfulness Data	66
4.1	Introduction	67
4.2	Optimism and Placebo	68
4.2.1	Factorial Structure of the LOT-R	70
4.3	Mindfulness	70
4.3.1	Health benefits of Mindfulness	71
4.4	Aims of the Research	72
4.5	Methods	72
4.5.1	Participants	73
4.5.2	Measures	73
4.5.3	Sampling for this research	74
4.6	Analysis	74
4.7	Results	75
4.7.1	Descriptive Statistics	75
4.7.2	Psychometric Analyses	76
4.7.3	RAND MOS	76
4.8	Mindfulness Attention Awareness Scale	79
4.8.1	Factor Analyses, Sample One	79
4.8.2	Factor Analyses, Sample Two	81

4.8.3	IRT Analyses - MAAS	83
4.8.3.1	Sample One	83
4.8.3.2	Sample Two	85
4.8.4	Back testing	89
4.9	Life Orientation Test, Revised	89
4.9.1	Factor Analyses, Sample One	89
4.9.2	Factor Analyses, Sample Two	91
4.9.3	IRT Analyses, LOT-R	92
4.9.3.1	Sample One	92
4.9.3.2	Sample Two	93
4.9.4	Back testing - LOTR	96
4.10	Health, Optimism and Mindfulness Analysis	96
4.10.1	Sample One, Relationships between Health, Optimism and Mindfulness	98
4.10.2	Regression Model Predictions	99
4.10.3	Regression Analyses, Sample Two	99
4.10.4	Optimism	99
4.10.5	Mindfulness Regressions	100
4.11	Modeling the relationship between health, optimism and mindfulness	100
4.11.1	Optimism and Health	101
4.12	Sample Two Structural Equation Models	101
4.13	Discussion	102
4.13.1	Optimism and Health	102
4.13.2	Mindfulness	103
4.13.3	Psychometric Analysis	104
5	Treatment Credibility Questionnaire	105
5.1	Introduction	106
5.1.1	Current Expectancy Measurement	106
5.1.2	Beliefs About Treatment	107
5.1.3	Links to theory	108
5.1.4	Pain Treatments	108
5.1.5	Aims and Objectives	109
5.2	Methodology	110
5.3	Design	110
5.3.1	Sampling	110
5.3.2	Participants	110
5.3.3	Measures	111
5.3.4	Analysis	112
5.4	Study One - Results	113
5.4.1	Demographic Statistics	113
5.5	Psychometric Analyses	114
5.5.1	Reliability Analysis	114
5.5.2	Factor Analysis	115
5.6	Item Response Theory Analyses	116
5.7	Confirmatory Analyses	119
5.8	Conclusions & Next Steps	121
5.9	Study Two - Results	122
5.9.1	Results	123

5.9.1.1	Descriptive Statistics	123
5.10	Treatment Credibility Questionnaire, Version 2	126
5.10.1	Five Factor Solutions	127
5.10.2	Six Factor Solutions	127
5.10.3	Seven Factor Solutions	129
5.10.4	Confirmatory Factor Analyses	129
5.10.5	IRT Analyses	129
5.10.6	Conventional Items	130
5.10.7	Alternative Items	133
5.10.8	Testing the Models	133
5.10.9	Final Tests	134
5.11	Reducing the size of the TCQ	136
5.12	Relationships between Credibility Totals and Other Variables	137
5.12.1	Discussion	140
6	Development of Implicit Measures	142
6.1	Introduction	143
6.2	Methodology	143
6.2.1	Repertory Grids	143
6.2.2	Development of IAT	144
6.3	Development of Repertory Grid	144
6.3.1	Piloting of the Repertory Grid	145
6.4	Development of IAT	146
6.4.1	IAT Stimuli	146
6.5	Piloting the IAT Measures	147
6.5.1	Optimism IAT Pilot	147
6.5.2	TCQ IAT	147
6.6	Piloting the Experimental Procedure	147
6.7	Discussion	152
7	Measuring Placebo by Multiple Methods	153
7.1	Introduction	154
7.1.1	Theoretical Models	155
7.1.2	Placebo Effects and Implicit Measures	155
7.1.3	Placebo effects and Deception	156
7.1.4	Explicit Measures and Placebo	157
7.1.5	Aims and Objectives	158
7.2	Methodology	158
7.2.1	Experimental Procedure	158
7.2.1.1	Recruitment	158
7.2.1.2	Measures	158
7.2.1.3	Procedure	159
7.2.2	Analysis of Experimental Data	159
7.2.2.1	Analysis of Reaction Time Data	160
7.2.2.2	Analysis of Implicit and Explicit Measures	161
7.2.2.3	Structural Equation Models	161
7.3	Results	162
7.3.1	Data Validation	162
7.3.2	Randomisation Checks	163
7.3.3	Analysis of IAT data	163

7.3.4	Explicit Measures	165
7.3.5	Relationships between Experimental Samples and Survey Samples	165
7.3.6	Implicit-Explicit Relationships	168
7.3.7	Primary Analyses	169
7.3.8	Predicting Placebo Response	171
7.3.9	Pain Ratings	174
7.3.10	Examining the Effect of Pain on GSR	176
7.3.11	Testing the Theoretical Models	180
7.3.12	Relationships between Explicit and Implicit Measures	180
7.3.13	Theoretical Models	181
7.4	Discussion	183
7.4.1	Relationship between Implicit Measures and Placebo	183
7.4.2	Modelling the Placebo Response with SEM	184
7.4.3	Physiological Analyses	184
7.4.4	Priming and Placebo Response	185
7.4.5	Scoring Methods	185
7.4.6	Conclusions	186
8	General Discussion and Conclusions	188
8.1	General Discussion	189
8.2	Health, Optimism and Mindfulness	191
8.3	Explicit Expectancy Measure	192
8.4	Repertory Grids and Experimental Pilot	193
8.5	Experimental Research	194
8.6	Strengths and Limitations	194
8.7	Overall Contributions	196
8.7.1	Implications for Clinical Research & Practice	196
8.7.2	Implications for Experimental Research	197
8.7.3	Methodological Contributions	197
8.8	Further Research	198
8.9	Conclusions	198
A	Appendix 1: Supporting Tables and Figures	201
A.1	Health, Optimism and Mindfulness Data Appendices	202
A.1.1	RAND MOS Psychometric Analyses	202
A.1.2	MAAS Psychometric Analyses	203
A.1.2.1	MAAS Three Factor Solution	203
A.1.3	LOTR Psychometric Analyses	204
A.2	TCQ Thesis	204
A.2.1	TCQ Factor Solutions	204
A.2.2	TCQ IRT Analyses	205
B	Appendix II: Qualitative Analysis	209
B.1	Introduction	210
B.2	Analysis of Health Construct Interviews	210
B.2.1	Methodology	210
B.2.2	Analysis	211
B.3	Methodology	211
B.3.1	Thematic Analysis	211
B.3.2	Participants	212

B.3.2.1	Alternative/Complementary Practitioners	212
B.3.2.2	Doctors	212
B.3.2.3	Students	212
B.3.3	Overall theme	212
B.4	Results	213
B.4.1	Health as balance	213
B.4.2	Energy Model of Health	214
B.4.3	Emotional Health	216
B.4.4	Personal Effort	218
B.4.5	Body Knowledge of Health	219
B.5	Construction of Doctors	220
B.6	Diet	226
B.6.0.1	Food as medicine	226
B.6.1	Diet as Fashion	227
B.6.2	Diet as Place Based	228
B.6.3	Nutrition	228
B.7	Models of Health	231
B.7.1	Health as balance	231
B.7.2	Health As Harmony	233
B.7.3	Health as Social Issue	234
B.7.4	Health is not a Problem	235
B.7.5	Health is not a priority	236
B.7.6	Linking Body and Mind	237
B.7.7	Linking Levels	238
B.7.8	More Than Physical	239
B.7.9	Energy Model of Health	241
B.8	Personal Factors	244
B.8.1	Awareness	244
B.8.2	Construction of Illness	245
B.8.3	Different Needs for Different People	246
B.8.4	Expectations Differ Across Class	247
B.9	Models of Sickness	248
B.9.1	Sickness Part of Life	248
B.9.2	Sickness from Pathogens	249
B.9.3	Sickness as a Way of Life	249
B.9.4	Services focus on Illness	249
B.9.5	Ill-health as Normal	250
B.9.6	Differences in Immunity	251
B.10	Social Dimensions	251
B.10.1	Class Differences	251
B.10.2	Family dimension of illness	254
B.11	Conclusions	255
C	Appendix III: Health Rep Grids	257
C.1	Health Rep Grids and Sorts	258
C.1.1	Instructions	258
	References	261

List of Figures

4.1	Item Information Curves for MAAS One Parameter GRM (left) and Two Parameter GRM (right) (Sample One)	85
4.2	Item Information Curves for MAAS One Parameter GRM (left) and Two Parameter GRM (right) (Sample Two, Split B)	87
4.3	Item Information Curves for MAAS One Parameter GRM (left) and Two Parameter GRM (right) (Sample Two, Split C)	88
4.4	Item Information Curves, LOT-R, Sample Two, Split B for One Parameter GRM (left) and Two Parameter GRM (right)	94
5.1	Item Information Curves, TCQ Conventional Scale, One Parameter GRM (top) and Two Parameter GRM (bottom)	119
5.2	Item Information Curves, TCQ Alternative Scale, One Parameter GRM (top) and Two Parameter GRM (bottom)	120
5.3	Scatter-plot Matrix, TCQ 2. The upper triangle represents correlations between the variables, the lower triangle shows scatterplots between them, and the diagonal shows a histogram of the values. From top to bottom: Pilltot - Pill Credibility, Creamtot - Cream Credibility, Injtot - Injection Credibility, Acutot - Acupuncture Credibility, Homtot - Homeopathy credibility, Reitot - Reiki credibility, BMQ - BMQ Scores	124
5.4	Credibility Scores by College	125
5.5	Hierarchical Factor Structure (Omega) for TCQ Data	131
5.6	Item Information Curves, TCQ Conventional, Split 2, One Parameter GRM (top) and Two Parameter GRM (bottom)	134
5.7	Information Curves for Items One and Three for TCQ-Conventional, Split A	137
6.1	Histogram of Frequency of People or Qualities	145
6.2	Distribution of Optimism IAT Scores	148
6.3	Scatter-plot of Correlations between LOTR and Optimism IAT	149
6.4	Pairs Plot of Explicit Variables, Pilot Study. LOTR: Life Orientation Test, Revised. MAAS - Mindful Attention Awareness Scale. Pill - Pill Credibility. Cream - Cream Credibility. Inj - Injection Credibility. Acu - Acupuncture Credibility. Hom - Homeopathy Credibility. Rei - Reiki Credibility	150
6.5	Participant individual pain responses over time	151
7.1	Density Plot for Optimism (top) and Mindfulness Scores (bottom), Survey Sample (red), and Experimental Sample (blue)	166
7.2	Density Plot for Pill Credibility (top) and Cream Credibility (bottom), Survey Sample (red), and Experimental Sample (blue)	167
7.3	Injection Credibility Density Plot, Survey Sample (sample two) (blue), Experimental Sample (red)	168

7.4	Scatterplots and Correlations for Implicit and Explicit Measures. Upper triangle shows the correlations, lower triangle shows the scatterplots, Diagonal shows the distributions. From left top: LOTR - Life Orientation Test, Revised. MAAS- Mindful Attention Absorption Scale. Pill - Pill Credibility. Cream - Cream Credibility. Inj - Injection Credibility. Acu -Acupuncture Credibility. Hom - Homeopathy Credibility. Rei - Reiki Credibility. OptIAT - Mean Scored D Measure, Optimism IAT. TCQIAT - Mean Scored D Measures, Treatment Credibility IAT . . .	170
7.5	Count of Placebo Response in Primed and Non-Primed Conditions . .	171
7.6	Pain Ratings of Participants by Response to Placebo Across Time. Straight line is a Loess smoother, the jagged line represents the actual pain levels	172
7.7	Plot of Pain Responses by Condition Over Time	175
7.8	Median Pain Ratings Over Time by Condition	176
7.9	GSR over Time by Response to Placebo	177
7.10	GSR by Condition over Time	178
7.11	Mean GSR Levels by Condition over Time	179
7.12	Direct Placebo Response Model (Model 1)	181
7.13	Placebo SEM Kirsch with implicit measures	182

List of Tables

4.1	Gender Breakdown of Participants by Collection Method	73
4.2	Summary Statistics for Health Scales, Mindfulness and Optimism, Sample One	75
4.3	Summary Statistics for Health Scales, Mindfulness and Optimism, Sample Two	75
4.4	Factor Loadings Eight Factor Solution, RAND MOS, Sample One	77
4.5	Factor Correlations, Eight Factor Solution RAND MOS, Sample One	79
4.6	SEM Comparison for RAND MOS Factor Solutions, Sample One	79
4.7	Model Comparison for RAND MOS data using Models from Sample 1 on Sample 2	79
4.8	MAAS Item Content	80
4.9	Factor Loadings, One Factor Solution, MAAS, Sample One	80
4.10	Comparison of One and Three Factor Models, Sample One	81
4.11	Comparison of Sample One MAAS Factor Models on a subset of Sample Two Data (Split A)	81
4.12	One Factor Solution Averaged over Splits B and C, MAAS	82
4.13	Comparison of Factor Structures for MAAS 2B Solutions, Tested on Split B	83
4.14	Comparison of Factor Structures for MAAS 2C Solutions, Tested on Split C	83
4.15	Coefficient Estimates for MAAS One Parameter Graded Response Model, Sample One	84
4.16	Performance of MAAS One and Two Parameter GRM's on Unseen Data (Sample Two, Split A)	84
4.17	Coefficient Estimates for MAAS One Parameter Graded Response Model, Split B	86
4.18	Performance of One and Two Parameter Graded Response Models for MAAS on Unseen Data (Splits A and C)	86
4.19	Coefficient Estimates for MAAS One Parameter Graded Response Model, Split C	86
4.20	Performance of MAAS One and Two Parameter Graded Response Models on Unseen Data (Splits A and B)	88
4.21	Full and Reduced Models for MAAS Tested on Sample One Data	89
4.22	Backtesting of MAAS IRT Models on Sample One Data	89
4.23	Life Orientation Test, Revised Item Content	90
4.24	Factor Loadings, One Factor Solution, LOT-R, Sample One	90
4.25	Factor Loadings, Two Factor Solution, LOT-R, Sample One	90
4.26	Comparison of CFA for the LOT-R Sample One Models on Sample One	91
4.27	Comparison of CFA Results for LOTR Sample One, One and Two Factor Models on a Subset of Sample Two (Split A)	91
4.28	Average One Factor Solution for LOT-R Across Splits B and C	91
4.29	Average Two Factor Solution for LOT-R Across Splits B and C	92
4.30	Comparison of LOT-R Split B Factor Solutions, Tested on Split C	92
4.31	Comparison of LOT-R Split C Factor Solutions, Tested on Split B	92
4.32	Coefficient Estimates for LOTR Two Parameter Graded Response Model	93
4.33	Performance of LOTR One and Two Parameter GRM's on Unseen Data (Sample Two, Split A)	93

4.34	Coefficient Estimates for One Parameter Graded Response Model, LOT-R, Split B	94
4.35	Performance of One and Two Parameter Graded Response Models on Unseen Data (Splits A and C)	95
4.36	Coefficient Estimates for LOT-R One Parameter Graded Response Model, Split C	95
4.37	Performance of LOT-R Split C One and Two Parameter Graded Response Models on Unseen Data	96
4.38	Performance of Full and Reduced Model for LOT-R on Sample One Data	96
4.39	Performance of LOT-R Split C One and Two Parameter Graded Response Models on Unseen Data	96
4.40	Spearman Correlations between Scale Totals across both Samples	97
4.41	Coefficients for General Health Regression Model with predictors chosen by Stepwise Selection on unseen data, Sample One	98
4.42	Coefficients for Stepwise Selected Model on Test Data, General Health (Sample 2A)	99
4.43	Coefficients for Stepwise Selected Regression Model for Optimism	99
4.44	Coefficients for Stepwise Selected Regression on Mindfulness	100
4.45	Comparison of Different Models between Optimism and Other Variables, Sample One	101
4.46	Comparison of Models for Health, Sample Two	101
5.1	Summary Statistics, Total Credibility Scores	113
5.2	Item reliability Statistics for the TCQ 1. Overall alpha of the scale was 0.92, following reverse scoring of the Acupuncture Items. Reverse scored items are indicated by the minus symbol. Raw/standard alpha (Cronbach's alpha), G6: Guttman's Lambda6 (squared multiple correlation) S/N: Signal/Noise ratio, alpha_se: Standard error of alpha estimation	114
5.3	Four Factor Solution, TCQ 1 Oblimin rotation	116
5.4	Four Factor TCQ correlations	116
5.5	Six Factor Solution, TCQ 1, Oblimin Rotation	117
5.6	TCQ Six Factor Solution Correlations	117
5.7	CFA on TCQ Models, Sample One	117
5.8	Coefficients for the TCQ 1 Conventional Scale, One Parameter Graded Response Model	118
5.9	Coefficients for TCQ 1 Alternative Scale, One parameter Graded Response Model	118
5.10	Coefficient Estimates for TCQ 1 Alternative Scale, Two Parameter Graded Response Model	119
5.11	Comparison of Four and Six Factor Models for the TCQ1 on Sample Two	120
5.12	Comparison of One and Two Parameter Graded Response Models, TCQ Conv, Split 2A	121
5.13	TCQ2 Summary Statistics	123
5.14	Experience with painkilling treatments by Gender	125
5.15	Income Levels for Study two	126
5.16	Self Reported Health levels, Study Two	126
5.17	Average of Six Factor Solutions over Splits B, C and D	128
5.18	Average of Factor Correlations, Six Factor Solution, TCQ2 (B,C,D)	129
5.19	Average of Seven Factor Solutions over Splits B, C and D	130
5.20	Averaged Factor Correlations for the Seven Factor Solution, Sample Two	131

5.21	TCQ Models from Split C tested against Splits A, B and D	132
5.22	TCQ 2 Models built on Split D, tested against Splits A, B and C . . .	132
5.23	TCQ 2 Factor Solutions From Split B, Tested Against Splits A, C and D	132
5.24	Item Difficulty Thresholds for TCQ Conventional items averaged over Splits B, C and D	132
5.25	Item Difficulty Thresholds for Two Parameter GRM, TCQ Conv, Aver- aged over Splits B, C and D	133
5.26	Item Difficulty Thresholds for One Parameter GRM on the Alternative Items, averaged over Splits B, C and D	135
5.27	Average Error of Approximation for One and Two Parameter GRM's on the Conventional TCQ Across Splits B, C and D	135
5.28	Average Error of Approximation for One and Two Parameter GRM's on the TCQ Alternative items Across Splits B, C and D	135
5.29	TCQ2 Models from each split tested against the full dataset	135
5.30	Coefficients for Pill credibility Regression on Held Out Data	138
5.31	Coefficients for Cream credibility Regression on Held Out Data	138
5.32	Coefficients for Injection credibility Regression on Held Out Data . . .	138
5.33	Coefficients for Acupuncture credibility Regression on Held Out Data	139
5.34	Coefficients for Homeopathy credibility Regression on Held Out Data .	139
5.35	Coefficients for Reiki credibility Regression on Held Out Data	139
6.1	Frequency of Exemplars, Rep Grid Study	144
6.2	Gender of Pilot Participants	148
6.3	Predicting Placebo Response from TCQ IAT Scores	150
6.4	Predicting Placebo Response from OPT IAT Scores	151
6.5	Predicting Placebo Response from Explicit Credibility Scores	151
7.1	Summary of ANOVA of Treatment Credibility IAT Scores by Condition	163
7.2	Summary of ANOVA of Optimism IAT scores by Condition	163
7.3	Mean and SD of GSR by Condition at baseline (first 100 seconds) . .	163
7.4	Correlations between the blocks of the treatment credibility IAT (Kendalls tau.) All correlations are significant at the $p < 0.001$ level . .	164
7.5	Correlations between the blocks of the Optimism IAT (Kendalls tau). All correlations are significant at the $p < 0.001$ level	165
7.6	Model selected by Stepwise Selection over the training data	173
7.7	Coefficients of Stepwise Selected Predictors on Test Data	173
7.8	Odds Ratios for Coefficients in Test Data Model	173
7.9	Summary statistics for the bootstrapped p-values	174
7.10	Regression of Placebo Response on IRT Estimated Optimism Scores .	174
7.11	Regression of Conventional TCQ Summary Scores on Placebo Response	174
7.12	Regression of IRT Estimated Ability Scores on Placebo Response . . .	174
7.13	Logistic Regression for the Impact of Mean GSR measurements on placebo response	174
7.14	Median Recording Time (secs) by Condition and Placebo Response. Time is counted from the beginning of physiological monitoring (ap- proximately 5 minutes before application of painful stimuli)	176
7.15	Regression of Condition on Mean GSR	180
7.16	Fit Statistics for the SEM Models on the Relationship between Implicit and Explicit Expectancies	180
7.17	Fit Measures for Placebo Response Structural Equation Models	182

7.18	Coefficients for Kirsch Two Factor Model	183
A.1	Factor Loadings, RAND MOS Two Factor Solution, Sample One . . .	202
A.2	Factor Correlations, RAND MOS Two Factor Solution, Sample One .	203
A.3	Factor Loadings, Three Factor Solution, MAAS, Sample One	203
A.4	Factor Correlations, MAAS Three Factor Solution, Sample One	203
A.5	Factor Loadings, MAAS Five Factor Solution, Sample Two, Split B . .	204
A.6	Factor Loadings, MAAS Five Factor Solution, Sample Two, Split C .	204
A.7	Coefficient Estimates for MAAS Two Parameter Graded Response Model	205
A.8	Coefficient Estimates for MAAS Two Parameter Graded Response Model, Split B	205
A.9	Coefficient Estimates for MAAS, Two Parameter Graded Response Model, Split C	206
A.10	Two Factor Solution, LOT-R, Split B	206
A.11	Two Factor Solution, LOT-R, Split C	206
A.12	Coefficient Estimates for Two Parameter Graded Response Model, LOT- R, Split B	206
A.13	Coefficient Estimates for LOT-R, Two Parameter Graded Response Model	207
A.14	Average of Five Factor Solutions over Splits B and C	207
A.15	Coefficients for TCQ 1 Conventional Two Parameter Graded Response Model	208
C.1	List of People	260

I, Richard Patrick Morrisroe, certify that this thesis is my own work and I have not obtained a degree in this university or elsewhere on the basis of the work submitted in this thesis.

Richard Patrick Morrisroe

To my parents, without whom none of this would have been possible

Abstract

This thesis investigated the relationship of explicit (self-report), implicit (IAT) and physiological variables to the placebo effect. The thesis consisted of three main parts. The first collected background data and developed models for two constructs (Optimism and Mindfulness) associated with the placebo effect and implicit attitudes, respectively. The second part of the thesis consisted of the development of an explicit measure of treatment expectancies, and the development of two IATs, one for Optimism and the other for Treatment Credibility. The final portion of the thesis was an experimental study (N=111) which tested these new measures in a sample of healthy volunteers. The primary hypothesis of the thesis, that there would be a relationship between the placebo effect and implicit measures, was not supported. Major findings include an effect of semantic priming on placebo response mediated by condition (Deceptive versus Open Placebo), an unexpected negative relationship between Optimism and self-reported Health, and a physiological relationship between pain ratings and GSR data, which was also mediated by Condition in the experiment.

Acknowledgements

Firstly, I'd like to thank all of the people who supervised me over the course of this rather long journey. These are Dr Zelda Di Blasi, Dr Dylan Evans, Dr Jurek Kirakowski, Prof John Groeger and Dr Sean Hammond. All of you have helped immeasurably in making this thesis better, and for that I thank you.

Next, I'd like to thank all of my friends who stuck with me throughout this long and winding road, covering three cities, four jobs and two relationships. On that note, I'd like to thank Nicola Ziedler for being there for the first four years.

Additionally, I'd like to thank all of the people who contributed to the open source software that helped make this thesis.

I would also like to thank the examiners who made this a much better thesis.

Finally, I'd like to thank Elizabeth Ann Doyle for putting up with a difficult ending to a really long process.

Chapter 1

Introduction

1.1 Preface

The central aim of this thesis was to examine the measurement of treatment expectancies in the context of the placebo effect. This thesis used explicit (self-report), implicit and physiological variables to deliver greater understanding of these psychological constructs and their relationship with the placebo effect.

The more specific aims of this research were as follows:

1. To create an implicit measure to assess expectancies related to treatment and to the placebo response;
2. To develop a multifaceted method of assessing explicit expectancies which can be linked to the placebo response;
3. To compare the predictive power of implicitly and explicitly measured constructs as predictors of the placebo response;
4. To conduct one experimental study using a placebo analgesia paradigm to assess the relationship between implicit and explicit expectancies and the placebo response.

The thesis was structured into a number of parts, each of which will be discussed in turn. Following this, the overall aims and objectives of the thesis will be reviewed and set into context. Next, the specific contributions of this research to the literature will be elucidated. Finally, some avenues for future research will be described.

The first section of the thesis was the review of the literature around the placebo effect and implicit measures, along with any known individual-level predictors of the effect. This review found that placebo effects have been seen to be mediated by optimism (as operationalised by the Life Orientation Test, Revised (LOT-R) (Scheier,

Carver, & Bridges, 1994)) (Geers, Helfer, Kosbab, Weiland, & Landry, 2005; Morton, Watson, El-Deredy, & Jones, 2009) in both pain and non-pain paradigms. This review also linked the constructs of optimism, which can be defined as “generalised outcome expectancies around the future” (Carver, Scheier, & Segerstrom, 2010), with the construct of response expectancies of Kirsch (Kirsch, 1985, 1997), which are described as the “expectation of a non-volitional response”. These constructs both crucially rely on the participants’ perception of the likelihood of the event, and a secondary aim of the thesis developed which was to examine the usefulness of these as two separate rather than one combined predictor. The hypothesis was that these two constructs should be significantly correlated, and that one of them would mediate the impact of the other on the observed placebo response (Geers, Helfer, et al., 2005).

Another important potential moderator variable which came to light in the literature review was the construct of mindfulness, as operationalised by the Mindful Attention Absorption Scale (MAAS) (Brown & Ryan, 2003). This variable was found to moderate the correlations between explicit and implicit measures in an experience sampling study. Furthermore, implicit measures were found to be better predictors of spontaneous responses, while explicit measures were found to be better predictors of deliberate behaviour (Levesque & Brown, 2007), a pattern known as the double-dissociation effect and observed in IAT studies in a variety of domains (Asendorpf, Banse, & Miercke, 2002; Perugini, 2005; Grumm & von Collani, 2007; Steffens & König, 2006).

This research literature suggested that a good strategy would be to replicate the LOT-R and MAAS on a number of samples from which the eventual experimental sample would be drawn, both to develop better models for these constructs’ relationship (as they had not been concurrently examined when this work was carried out) and to tailor the measures towards the experimental sample, a strategy which was hoped to increase the precision of estimation on the overall sample. In this research, the RAND-MOS was used both to replicate the impact of these primary constructs on health, and to provide a weak link to health and treatment expectancies in preparation for the experimental research.

One large problem with the development of an implicit measure of treatment expectancies was that there exists no gold-standard measure of explicit treatment expectancies, despite them having been shown to materially affect the outcome of clinical trials (Linde et al., 2007; Bausell, Lixing, Bergman, Wen-Lin, & Berman, 2005; Benedetti, 2005). This was an issue both because typical methods of IAT development rely on such a self-report measure, and because without such a measure, the incremental improvement (if any) granted by an implicit measure would not be clear. Therefore, the development and testing of such a measure formed part of the work of this thesis.

Another fact which became extremely clear during the course of the literature review, is that current methods of both developing and validating implicit measures (or more specifically, IATs), rely extremely heavily on the existence of prior self-report instruments. For the placebo effect which was the focus of primary investigation, no such self-report measure existed, which meant that a new approach needed to be tested. A review of methods in this space revealed the work of George Kelly on personal construct theory (G. A. Kelly, 1991), and more specifically the repertory grid introduced in Book 1 of the aforementioned reference. It was believed that such an approach might prove useful in the current case, and might also prove useful with other constructs of interest in the future¹.

Therefore, in order to develop these stimuli, it was decided to do some qualitative research around health constructs with doctors, alternative therapists and students, to gain an appreciation for how they contextualised and understood issues around health, medicine and sickness. The results of this procedure, along with a parallel exercise asking participants to rate the five most important people or qualities of people related to health, would then be used to develop the Health Repertory Grid.

Finally, the development of the experiment was planned, and it was decided to use a placebo analgesia design, inducing pain with the sub-maximal tourniquet technique (Moore, Duncan, Scott, Gregg, & Ghia, 1979), as this form of pain induction has been shown in a meta-analysis to produce the largest effect sizes (Sauro & Greenberg, 2005).

The remainder of this chapter provides a short introduction to some of the major areas of focus of the thesis, and both the placebo and implicit measures are much more exhaustively described in Chapter 2.

1.2 Introduction to Placebo & Implicit Measures

1.2.1 Placebo

The placebo is a complex topic of study in that the term stands as a proxy for those elements of human health which are not determined by features specific to the treatment. Instead, the term placebo refers to the “non-specific” parts of the treatment (Grünbaum, 1981), those which are not attributable to specific biological or mechanical activities of the treatment, but are observed across varying treatments. In controlled experiments, the placebo effect is often defined as the response in the placebo group less the response in the no-treatment group. While this definition has problems (described in Section 2.2), it is good for an intuitive understanding of the construct.

¹I am indebted to Drs Jurek Kirakowski and Sean Hammond for originally suggesting this idea.

The placebo effect has been studied in and of itself for approximately 60 years since Beecher at least (Beecher, 1955). The research study of placebo arose as a result of the requirement for placebo controlled studies of medical interventions, and the first such trial was for the efficacy of streptomycin (Concato, Shah, & Horwitz, 2000). In the time since then, there have been many studies to test the relationships between individual level constructs and the phenomenon. For the most part, these attempts have been unsuccessful. A review of the problems with many of the earlier approaches is in Shapiro and Shapiro's book (Shapiro & Shapiro, 1997). The central idea of this research is that this lack of success in prediction is an artefact of the measures used.

More recently, there have been some intriguing studies that suggest that there may be some relationships between optimism and placebo response (Geers, Helfer, et al., 2005; Morton et al., 2009).

The placebo effect is rarely the focus in randomised controlled trials, and yet that is where much of the research which has broadened our understanding of the phenomenon has taken place. Despite years of research, there are few known individual level predictors of the effect. In fact, some researchers have argued that there is no such entity as a placebo responder (Kaptchuk, Kelley, Deykin, et al., 2008).

There is some confusion surrounding the definition and interpretation of placebo (c.f. Section 2.2.1). The term was retained throughout this research as it does form a useful overarching construct for this research on the interactions between mental and physical states in health. A definition suitable for the purposes of this thesis was the following:

a placebo is a treatment believed inert for the specific condition concerned which has an effect due to context.

The rationale behind this definition (and an explanation of context) is given in Chapter 2, in Section 2.2. Note that typically a “placebo” refers to the treatment, the “placebo effect” is the effect in a placebo treatment arm of some kind, and a “placebo response” refers to the impact of a placebo on a particular individual.

1.2.2 Treatment Expectancies

The current prevailing theory around what causes the placebo effect in humans is the response expectancy model of Kirsch (Kirsch, 1985, 1997), which suggests that placebo effects are directly mediated by these expectancies, which are defined as “the expectation of a non-volitional response”.

For Kirsch, these expectancies are regarded as the ultimate causative force behind the placebo response in participants and patients, in that all other factors are mediated by them. This theory was regarded as being confirmed in relation to other theories (especially the conditioning theory of Voudouris (Voudouris, Peck, & Coleman, 1985)) by a 1997 paper which showed that the large positive effect of surreptitious reduction of pain to induce conditioning could be abolished by informing participants that this was occurring (Montgomery & Kirsch, 1997).

Despite the importance of expectancies in the field at present, no standardised measure of expectancies exists, and indeed a recent study showed that of sixteen placebo trials, only two had an expectancy measure in common (Myers et al., 2008).

One of the largest problems with expectancy research at present within the field of placebo is that the measurement of expectancies tends to be quite superficial. Typically, one question on an 11 point scale is asked and the response to this question is taken as the participants expectancy for the treatment. This conflicts with evidence that expectancies are far more multi-dimensional than this approach suggests (Stone, Kerr, Jacobson, Conboy, & Kaptchuk, 2005). Indeed, one of the aims of this thesis was to address this problem by developing a better scale for the measurement of expectancies (Chapter 5).

Given the expectancy theory's superiority within the field at present, it is important to look more closely at what the term means. Some authors (Stewart-Williams & Podd, 2004) argue that expectancies are necessarily conscious, which is a position which seems improbable, given the lack of awareness that typically accompanies observed placebo effects, and the deception which appears endemic to the field of study (Miller, 2008; Miller & Kaptchuk, 2008), in that participants in experimental studies are deliberately lied to in order to induce a placebo effect.

Expectancy is a catch-all phrase, and while it appears to have applications in a wide variety of areas (Montgomery, David, DiLorenzo, & Schnur, 2007) the term is far too broad to focus research specifically. Indeed, a similar "conditioning" paradigm has been used to produce "placebo" effects on sensory reports (Sterzer, Frith, & Petrovic, 2008), suggesting that these expectancy related mechanisms may be useful predictors outside the study of the placebo effect.

The value of the expectancy framework is that it has provided both a common vocabulary and a common mechanism for the measurement of placebo responses. The very broadness of the construct allows for it to be used in a variety of situations, which has ensured its survival in the field. This is also the worst part of the definition, as its wide-ranging applicability coupled with its lack of falsifiable predictions has meant that it is regarded as a theory in the abstract, but research on the determinants and measurement of expectancies is not prominent in the field at present.

Another issue with expectancies and their convergent validity is the relationship between the response expectancies of Kirsch and the generalised outcome expectancies used in the study of optimism (Carver et al., 2010). Both of these constructs are similar enough that the use of having both of them needs to be determined, and this analysis was carried out as part of this research (Chapter 7). Kirsch also notes that self-efficacy (Bandura, 1977) and response expectancies correlate quite highly (Kirsch, 1985) and given this, the usefulness of a separate response expectancy framework needs to be examined, especially in the light of some of the findings discussed in Chapter 2 (c.f. (Geers, Helfer, et al., 2005; Hyland, Whalley, & Geraghty, 2007), and discussion in Chapter 4). Note that as expectancies were the primary focus of this research, discussion of other theories is postponed until the next chapter (see Section 2.3).

Given that the placebo seems to rely upon the belief of a participant that they have received a real treatment, it seems likely that measures which examine conscious beliefs are not the appropriate kind with which to predict its occurrence. It is the contention of this research that implicit measures may provide a better metric for the prediction of placebo response.

1.2.3 Moving Beyond Self-Report

Psychology as a science, and indeed the social sciences more generally, have a problem. They seek to understand the mind and behaviour of individuals in particular contexts and cultures. While behaviour is less problematic to observe scientifically, the observation of mind is fraught with problems. Almost all of the constructs of interest to psychologists (mind, love, experience) are not observable with the naked eye, and require interpretation in order to be understood.

In the case of many variables, psychologists measure what people think and believe from their answers to questions devised by the psychologist, normally referred to as self-report instruments.

This approach has obvious advantages, in that it is quick, cost-effective and can produce results of interest. However, as psychology has matured, a number of problems have become apparent with this approach (Nisbett & Wilson, 1977). The first problem is that, especially in controversial topics, people may attempt to conceal their true beliefs or attitudes. The second, somewhat more philosophical problem, is that people may be unaware of their true beliefs, or at least may profess to believe one thing while behaving in a manner consistent with a belief in another.

The first of these problems is known as social desirability (Egloff & Schmukle, 2003), and self report scales that measure this construct have been developed (Giebel & Groeben, 2008). The second problem, first noted by Freud, is that of unconscious (or

implicit) influences (Hofmann & Schmitt, 2008). While the system built by Freud no longer forms part of the framework of modern psychology, the contradictions between reports of experience and behaviour remain, and are still relevant to the aims of psychology.

1.2.4 Older non-self report methods

Some methods have been developed to get around this problem. One of the first techniques used for this purpose was that of free association, where a client was asked to respond to a stimulus word or picture with the first word that came to mind, without censoring the experience (Hofmann & Schmitt, 2008). These associations could then be used by the therapist to gain access to material which the client did not consciously report being aware of.

Another method which was used was that of Rorschach ink-blots, where ambiguous ink blots are shown to the client, who interprets them. This technique can also provide insight into the mind of the client, but again this requires interpretation on the part of the therapist. It is this interpretation process that causes these procedures to lack scientific validity in the eyes of many, as what one therapist understands by the clients words may differ completely from what another therapist takes from the same material.

These approaches were abandoned following the rise of behaviourism as the dominant approach within academic psychology, and further eclipsed by the notions of Karl Popper regarding falsifiability as a criterion for scientific theory (Popper, 1954). It was argued that since unconscious influences could be used to explain any criticism of the theory (and indeed, Freud was prone to doing this) then the theory was not truly scientific. The development of psychometric theory also played a role in the decline of interest in such instruments, as these methods seemed to produce reliable and valid data and scores could be corrected for impression management and social desirability biases by statistical techniques.

1.2.5 Modern Indirect Measures

In recent years, however, there has been a resurgence of interest in such techniques. This resurgence grew out of the work on implicit memory and learning, where participants would consciously deny awareness of some piece of information while their behaviour seemed to show signs of this knowledge.

This phenomenon can be seen in experiments like word completion tasks. If participants are given a list of words to memorise, and then distracted by another task, followed by the word completion task, they tend to far more frequently complete

the word fragments with the words on the previous list which was to be memorised. However, they will typically deny this influence on their responding if asked (Wittenbrink & Schwarz, 2007a).

These approaches, allied with the continuing failure of self reported attitudes to predict behaviour as accurately as might be desired caused some researchers to look for another way to measure these constructs (Greenwald & Banaji, 1995). The result of these investigations was the Implicit Association Test, or IAT for short (Greenwald, McGhee, & Schwartz, 1998). The IAT is a reaction time measure which makes inferences about attitudes from the time which it takes participants to categorise words or pictures on a related theme into one of two categories. The difference between the participants response latencies is taken to indicate the relative strength of associations, and this difference is known as an IAT effect.

1.2.6 Introduction to the Implicit Association Test (IAT)

The IAT was developed by Greenwald *et al.*, and he suggested that because of its design, it might be more resistant to social desirability influences and demand characteristics (Greenwald et al., 1998).

Social desirability tendencies would lead people to deny prejudicial behaviour in self report instruments, while as a reaction time measure, the IAT is less easy to fake. Demand characteristics result when participants in an experiment give the answers a researcher wants, rather than their real beliefs or attitudes. Again, it seems intuitively harder to do this within an IAT methodology as it would require tremendous and consistent control of reaction times to stimuli².

1.2.7 Description of the Procedure

The (IAT) is a computer administered procedure which purports to measure implicit associations not directly accessible to consciousness (Greenwald et al., 1998). The test was developed as a result of mounting evidence for learning without awareness in human participants, as discussed above.

This research was reviewed by Greenwald & Banaji (Greenwald & Banaji, 1995), where they introduced a distinction between direct and indirect measures of social cognition. They referred to self report instruments as direct measures, and to such techniques as semantic priming as indirect measures. Semantic priming is the tendency for participants in experiments to give answers to ambiguous tasks similar to ones they have recently observed in their environment (Wittenbrink & Schwarz, 2007b). They defined implicit associations in the following manner as *the unidentified*

²though not impossible (De Houwer, Beckers, & Moors, 2007)

or inaccurately identified traces of past experience, and this definition implied that self report measures were not the best tools with which to assess these associations.

The procedure works as follows. Firstly participants sit down in front of a computer and are assigned two keys (typically the “e” and “i” keys) to respond to each word or image presented, which fall into one of two categories. The metric examined is reaction time, and an assumption of the method is that categories which are more strongly associated will be easier to combine than those which are less associated.

The participant is asked to classify words into either pleasant or unpleasant categories as they appear on the screen at the front of the computer, for example love, hate, good and bad are words typically used in this part of the procedure. In the most well known IAT, the Race-IAT, the procedure is as follows:

Firstly, participants are asked to categorise faces into as either being “Black” or “White”³. Following this, the two categories are combined, with one key being pressed for White or Pleasant and another being pressed for Black or Unpleasant. Then, the labels are reversed, and the participant categorises White faces with Unpleasant and Black faces with Pleasant. The original authors described the White + Pleasant trials as congruent, and the Black + Pleasant/White + Unpleasant trials as incongruent. These response times are summed and averaged for each participant for each of the critical blocks, and the two block scores (Incongruent - Congruent) are subtracted from one another to produce a difference score which is referred to as an IAT effect.

The procedure is not limited to assessment of racial attitudes, and has been applied far more widely (Craeynest, Crombez, Koster, Haerens, & De Bourdeaudhuij, 2008; Greenwald, Poehlman, Uhlmann, & Banaji, 2009; Schmukle, Back, & Egloff, 2008; Walker & Schimmack, 2008). A general schema for the process follows.

Firstly participants classify words as either belonging to Category X or Y, where X and Y are positively (love, flowers etc) or negatively (hate, insects etc) associated words or images. Then they classify faces or words as either belonging to Group A or Group B, where the words are often descriptive of groups of people. In the next step, these two associations are combined, with one key being a response for A and X and the other key being used for responses of B and Y.

In the fourth step, the keys for pleasant and unpleasant are reversed, and in the final step the two dimensions are combined in the opposite manner (A and Y or B and X). In practice, only the 3rd and the 5th steps are analysed, and the difference between mean response latencies on the different combination tasks is known as an IAT effect (Greenwald et al., 1998).

In essence, any differences in reaction time in the combination of the two categories

³Following some research around this, the categories were later re-named to African-American and Caucasian-American

are assumed to be due to underlying differences between the relative associations of the concepts. The authors claim that the use of difference scores allows them to prevent issues of processing and response speed variability across individuals from distorting the results. However, this assumption has been questioned (Blanton, Jaccard, Gonzales, & Christie, 2006), who claimed that processing speed was a major moderator of the observed variance in scores across the population studied. Controls were taken in the analysis of IAT scores to ensure that such effects did not impact the results (c.f. Chapter 7).

1.3 Structure of this Thesis

This thesis follows the following structure:

1. Chapter 2 reviews the literature surrounding the placebo response and explicit and implicit expectancies;
2. Chapter 3 describes a model for how the placebo effect and implicit and explicit expectancies inter-relate, and describes the approaches taken in this thesis for both the construction of measures and the testing of these models;
3. Chapter 4 reports the analysis of self report variables which have been associated with placebo effect (optimism) as well as with response to implicit measures (mindfulness), and the construction of tailored scales for the population at hand;
4. Chapter 5 details the development and validation of the Treatment Credibility Questionnaire, the self report method used to assess expectancies;
5. Chapter 6 describes the approach taken to the development of the stimuli for the implicit association tests, using multiple methods;
6. Chapter 7 reports upon the experimental portion of the research, and models these results incorporating the models developed for the self report measures in Chapters 4 and 5;
7. Finally, Chapter 8 summarises what has been learned from this thesis, and gives directions towards future research.
8. The Appendices supply supporting tables (A for the quantitative research) and qualitative analysis (B) which was not core to the rest of the thesis.

Chapter 2

Literature Review

2.1 Introduction to this review

This review has three sections. Firstly, the placebo effect is reviewed in greater detail than in the previous chapter, with a particular focus on the measurement of placebo expectancies and other factors which are thought to influence this effect.

Next, the implicit measure used in this research is reviewed (the Implicit Association Test) (Greenwald et al., 1998) with a particular focus on its psychometric properties and its relationship to explicit measures of the same construct.

Finally, the evidence which links these two areas together is reviewed, and the major aims of this thesis are defined more completely.

2.2 The Concept of the Placebo

As discussed in Chapter 1, the placebo effect has a long history in medicine, and some researchers believe that almost every treatment developed before the 20th Century may have relied primarily on this effect for their curative properties (Shapiro & Shapiro, 1997; Macedo, Farrao, & Baatos, 2003). Despite this, the concept is not particularly well defined, and different researchers use the same phrase to mean different things (Ernst & Resch, 1995; Hrobjartsson, 1996). A definition suitable for this thesis was given in Chapter 1 and the reasoning behind this definition is outlined in this section.

2.2.1 History of the Concept

Placebos have a long and colourful past, some elements of which are still relevant today. The word itself is taken from the Latin for "I will please" and referred originally to the cries of paid mourners at funeral ceremonies (Macedo et al., 2003). Over time, it became associated with any medicine which was given more to please a patient than to actually relieve the symptoms (Kaptchuk, 1998). Placebos were considered ethically acceptable, if somewhat dubious, for many years. However, as more effective and tailored treatments were developed, the use of placebos declined (Macedo et al., 2003). Following the development of randomised double-blind trials, the placebo increased in importance once more (Kaptchuk, 1998), as it became a required part of the process for establishing the efficacy of all new drugs.

The first serious attempts at studying of the placebo as an effect in itself were begun by Beecher's article "The Powerful Placebo" which argued that placebos were useful in many differing kinds of ailments and diseases, and that their effectiveness did not depend on the intelligence of the patient, as had previously been supposed (Beecher,

1955). This article aroused interest in the placebo, and is generally regarded as the first paper which focused on the effect in its own right, rather than as a device to disguise a doctor's lack of effective treatments (Beecher, 1955; Kaptchuk, 1998).

Beecher made many strong claims in his 1955 article, including an assertion that placebos were effective 35% of the time. Later research has demonstrated that many of these claims around the efficacy of placebo did not account for all potential confounding factors (Kienle & Kiene, 1998). These researchers re-analysed the studies reported on by Beecher, and found that much of the reported improvement could have been due to natural history of the disease and regression to the mean. Additionally, later research has shown that placebo effects can be much more variable than 35% (Turner & Deyo, 1994), especially when the median rather than the mean placebo response is considered (McQuay, Carroll, & Moore, 1996).

Much of the lack of clarity found within the field (Macedo et al., 2003; Kaptchuk, 1998) results from the use of placebo in two contradictory situations. In the first, that of the randomised controlled trial (RCT)¹, placebos are a control for all effects of treatment not related to the substance or procedure under test (Vickers & de Craen, 2000). In this setting, the aim is that they should be minimised so as the impact of the active treatment can be assessed.

In the second setting, that of clinical practice, the placebo is imbued with all the authority of medicine and utilised in order to effect changes that may result from mindset or to placate a troublesome patient (Bootzin, 2003; Sherman & Hickner, 2008). Both Macedo and Kaptchuk argue that both of these approaches give too much power to the concept and contribute to the confusion surrounding its definition.

Another confusing factor within placebo research results from the terms “specific” and “non-specific” effects, introduced by Grunbaum in 1981 (Grünbaum, 1981). These particularly confusing terms have led to much agonising and debate over the years. Some have even suggested that these terms should be abandoned (Caspi & Bootzin, 2002).

The specific parts of a treatment are typically defined as the biologically or theoretically effective agent, while the non-specific factors are all other parts of the treatment. Price and Benedetti argue that there are three sources of non-specific effects — the patient, the provider and the relationship between them (Price, Finniss, & Benedetti, 2008), a distinction also made by other authors (Finniss & Benedetti, 2005; DiBlasi, Kaptchuk, Weinman, & Kleijnen, 2002). It is important to note that more of the conceptual confusion arises from the differences between a placebo and a placebo effect. In brief, the placebo is the device or treatment given to a person, while the placebo effect is the outcome of this intervention. The term placebo response is

¹typically consisting of an experimental and control group with a double-blind design

often used to describe the impact of such a treatment on a particular participant in a clinical or experimental study, as described in Section 1.2.1.

2.2.2 Placebos in Randomised Trials

Placebos are most well known for their use in a clinical trial setting. In a typical randomised trial, neither patients nor physicians know whether a particular person receives drug or placebo. If there is any significant changes in the placebo group (with respect to the primary outcome measure of the study), even in the absence of medication, this is called a placebo effect, and the inert procedure or pill used is a placebo.² More technically, any mean improvement in the control group can be classified as a placebo effect. However this definition is not entirely accurate, as will be seen below.

The problem with this definition is that the effects in a placebo arm of a clinical trial result from a combination of the placebo effect and other factors, such as regression to the mean (Morton & Torgerson, 2003), demand characteristics of clinical trials (Hrobjartsson & Gotzsche, 2001; Weber & Cook, 1972) and other factors such as the natural history of the disease. Regression to the mean is the tendency for an extreme score measured at time 1 to be closer to the mean at time 2. This tendency is a property of all measuring tools which are not perfectly reliable (Morton & Torgerson, 2003)

Demand characteristics (Fernandez & Turk, 1994; Weber & Cook, 1972) refer to the tendency for participants in research to give the answer which they believe the researcher wishes to hear. This can result in an over-exaggeration of symptoms at the first assessment and a minimising tendency for the same symptoms at the end of the study (Vase, Robinson, Verne, & Price, 2005).

Without a no-treatment control group the effects of these confounders cannot be separated from the true placebo response. The no-treatment group serves this purpose as factors such as regression to the mean and natural history should apply equally to both the placebo and the no-treatment group. The definition of the placebo typically used is:

the placebo response is the response to treatment in the placebo group less the response to treatment in the no-treatment group.

A more precise definition in the context of clinical trials is given by Knipschild et al (Knipschild & Arntz, 2005).

²Often, if the effect is negative in terms of treatment outcome, the response may be called a nocebo response

the placebo effect ... [is] ... the difference in effect between the placebo group and the spontaneous course in a randomised clinical trial

This definition relies upon an understanding of the spontaneous course of an illness in a controlled trial, which can be operationalised as the progress of the no-treatment control group. It is somewhat limiting, but is clearly defined within a particular setting.

2.2.3 Problems with the clinical definitions of placebo

Within the confines of the clinical trial, the two definitions above are workable definitions of the placebo effect. The important part of a clinical trial is the test of the active medication, and the placebo is important only insofar as it relates to the testing of this medication.

However, clinical trials are not the only context where placebos are administered, and in other situations these definitions run into problems. Consider some recent work of Oken et al (Oken et al., 2008). In this experiment, healthy seniors were administered placebos which they were told would improve cognitive performance. There were two active groups (given different instructions) and one no-treatment group. The participants were tested for cognitive abilities at the beginning, middle and end of the placebo treatment. The seniors given the placebo pills showed significant increases in cognitive ability over the course of the study, and many were disappointed when debriefed and were told that they had been given placebos.

While the definitions given above can, at a stretch, account for these results, it does indicate a need to more carefully define the concept of placebo. Whether or not one accepts the clinical trial definition depends crucially on one's definition of treatment.

For many, this is some device or procedure that restores the organism to optimal function. Alternatively, this view could be described as believing that treatment restores homeostasis to the organism. Indeed, the Oxford English Dictionary agrees with this definition describing treatment as "*medical care for an illness or injury*" (Dictionary, 2010).

It is relatively easy to see how placebos can reduce pain, but is more difficult to see how such pills can improve cognitive performance. One can argue that decreases in pain result from a response bias (Allan & Siegel, 2002), but the measurements of cognitive performance in the Oken et al study were not subject to these kinds of bias.

Another study where placebo effects were demonstrated in a non clinical setting was that of Crum (Crum & Langer, 2007). In this study attendants in hotels were cluster randomised (using the hotel as the unit of sampling) and half of the attendants were informed of how many calories they burned by engaging in their roles as hotel

attendants. One month later, the informed group had lost significantly more weight and had improved on both self report and externally measured dimensions related to weight and health. Again, this effect is difficult to conceptualise in terms of treatment. Crum, in this paper defined the placebo effect as

any effect of treatment which is mediated more by the participants beliefs and expectancies rather than the physiological actions of the treatment.

This definition is more widely applicable than our first definition above, but at the cost of introducing two new terms which are not clearly defined, namely beliefs and expectancies. While the second of these terms has a number of specific meanings in psychological thought (for example, process, outcome and response expectancies) (Bandura, 1977; Kirsch, 1985), belief has not been so precisely defined. This definition also suffers from the issues surrounding the definition of treatment that were noted above, in Section 2.2.2.

Another definition (which avoids the treatment definition problem) was quoted from Ross & Olson, 1981 by Flaten and colleagues (Flaten, Simonsen, & Olsen, 1999) This definition is as follows:

A placebo or nocebo may be defined as an inactive substance or a procedure that is administered with suggestions that it will modify a symptom or sensation

This definition avoids the problem of defining treatment, and also avoids the use of the terms beliefs and expectancies. However, it can be seen from the definition that this is achieved only at the cost of introducing suggestion as a possible mediator of the placebo effect. On the face of it, this is perhaps not a term without merit. In studies of hypnosis, suggestion is commonly regarded as the driver of the observed effects (Kirsch, 1994) and hypnosis has been proposed as an ethical method to induce placebo responses in individuals (Raz, 2007). This definition is perhaps the best of those that have been examined so far, but it does require us to limit ourselves in the study of placebo to symptoms and sensations. Both the Crum & Langer study and the Oken *et al* study indicate that placebos have a wider sphere of effect.

Another definition of the placebo effects, proposed by Price *et al* (Price et al., 2008) is that

a placebo is any substance or procedure that simulates a treatment.

This definition fails to resolve our problem discussed above with relation to cognitive abilities. Kienle & Keine give two definitions of placebo in their critique of Beecher:

Placebo is defined in two separate ways, firstly as the imitation of a therapy, and secondly as any self-healing effect.

The first definition of imitation of a therapy is extremely similar to the Price definition quoted above, and the second, while it sounds plausible, is far too vague to be of any use in research.

Another definition (DiBlasi, Harkness, Ernst, Georgiou, & Kleijnen, 2001), does resolve the problems encountered with the Price & Benedetti definitions above. This definition is that

placebos are inert substances that have an effect due to context

This definition has some good points, in that it allows for placebo effects in any areas in which they are found, it allows for patient-provider effects and it does not prejudge the causes of the effect.

One major difficulty which can be observed with this definition is that it does not account for active placebos, where a substance which is a medication in one context is administered for a condition where it is not expected to have any pharmacological effect (Kirsch & Sapirstein, 1998) .

These active placebos would contradict the definition of placebo as inert, and yet the research demonstrates that these active placebos can be as effective as the regular inert pill (Flaten et al., 2004) and sometimes more effective (Kirsch, Moore, Scoboria, & Nicholls, 2002).

A similar definition (in the context of clinical trials) is given by Knipschild et al (Knipschild & Arntz, 2005) who say that the placebo effect

... is the effect of co-interventions in a treatment study connected to the doctor-patient relationship.

Again, this definition is quite precisely operationalised, but it assumes that the active ingredient of placebo is the relationship between provider and patient, and this has not been demonstrated to be true as some research into minimising contact between providers and patients has still demonstrated a placebo effect (Hyland, Geraghty, Joy,

& Turner, 2006), indicating that the patient-provider relationship cannot account for the entirety of the effect.

2.2.4 Definition for this thesis

Perhaps the most useful definition for the purposes of this thesis is the definition of Di Blasi et al (DiBlasi et al., 2001), if we take account of the issue arising from active placebos. The definition would work better if we removed the word “inert” and replaced with “believed inert for the specific condition concerned”, so that it reads

a placebo is a treatment believed inert for the specific condition concerned which has an effect due to context.

This definition would need to be supplemented with a more precise definition of context. One attempt at this would be that context is:

the internal states, external environment and relationship of the individual to these states, environment and other organisms in their presence.

2.3 Theories of Placebo Effects

Following the discussion of the definitions of placebo and some problems arising from the different contexts within which it has been used and defined, the next step in this review is to examine the theories which have been proposed to account for this phenomenon.

There are a few major theories, each of these will be described in turn and examined for the ways in which they account for the effects, and those features of the effect which they fail to explain (or reject as being illegitimate). A useful run-down of all of these theories appeared in the past few years (Stewart-Williams, 2004) and that review has helped to inform the arguments presented here.

2.3.1 Conditioning

The first major theory which attempted to account for placebo effects was that of conditioning. Building off the demonstration of placebo effects in non-human mammals (Herrnstein, 1962) the conditioning theory argued that placebo effects resulted from the learned association between a contingency in the environment (the

doctor, pill or medical setting) and healing. This contingency leads to the activation of healing mechanisms based on previous experience with the pill (Vodouris, Peck, & Coleman, 1989; Voudouris et al., 1985).

The conditioning theory has a number of advantages. Firstly, it can account for placebo effects in all mammals, as all seem to be capable of learning through reinforcement. Secondly, it is parsimonious, as it allows us to explain the placebo phenomenon without invoking any new processes or mechanisms. Thirdly, it appears to account for much of the observed effects.

One major problem for the conditioning theory is that it cannot account for placebo effects from a product which a participant has not experienced before. Given the nature of clinical trials, this rules out conditioning as a sole explanation for all placebo effects. One could assume that the observed placebo effects in clinical trials result from generalised associations with medical treatments more generally (Pearce, 1987). This does retain aspects of the theory, and provides a testable hypothesis, which is the following - to the extent that there are commonalities between the learning environment and the clinical trial environment, placebo responses will be observed.

The conditioning theory was strengthened by the demonstration that if pain is induced by means of a particular method, and the level of pain is surreptitiously reduced while the placebo is given, then participants typically show much stronger placebo responses (Voudouris et al., 1985; Colloca & Benedetti, 2006). Indeed, even though the conditioning theory has fallen from favour within the field, this methodology is still used to increase the size of observed placebo effects.

There do appear to be some placebo responses which are totally mediated by conditioning (such as hormone secretions) (Amanzio & Benedetti, 1999), but not all of them are (Benedetti, Pollo, et al., 2003). Experimental research has elucidated some of the connections here, in that motor movement in participants with Parkinson's disease and pain can be modulated by expectancies while changes in hormone secretion appear to be modulated by conditioning exclusively (Benedetti, Pollo, et al., 2003).

Strangely enough, even though conditioning appears to induce stronger placebo responses than does expectancies, nocebo suggestions can completely reverse the effects of positive placebo conditioning (Benedetti, 2008)³. Additionally, in one study, conditioned participants showed a greater placebo response when given neutral, rather than positive, instructions (Klinger, Soost, Flor, & Worm, 2007).

Some authors have argued (Stewart-Williams, 2004) that conditioning is not a theory of what causes placebo effects, but rather a mechanism through which other variables, such as expectancies, exert their influence. This is an interesting idea, and has some

³Possibly this occurs because of the impact of salience asymmetry

merit, but it does suggest that all placebo effects are mediated by expectancies. This does not match up with other research showing that dependent on the context of the study, other variables may be significantly related while expectancies are not (Geers, Weiland, Kosbab, Landry, & Helfer, 2005; Hyland et al., 2006) (and c.f. Section 4.2).

2.3.2 Expectancies

The competing theory to conditioning for the past few decades was the expectancy theory, as proposed by Kirsch (Kirsch, 1985), which was briefly reviewed in Section 1.2.2. Kirsch coined the term “response expectancy” to describe what he called “the expectation of a non-volitional response”. In a ten year review (Kirsch, 1997) suggests that this theory has applications in hypnosis and placebo effects. Recent research has shown that expectancies can also modulate sensory experience (Sterzer et al., 2008).

This theory competed with the conditioning theory for over a decade, but the issue was mostly resolved by a 1997 paper (Montgomery & Kirsch, 1997), which pitted the expectancy and conditioning explanations against one another. This study used the conditioning manipulation devised by Voudouris (Voudouris et al., 1985) where the painful stimulus is reduced after application of a placebo cream to increase the size of the placebo effect.

One group was told of the pain reduction, while the other was not. The group who were told showed no enhanced placebo response, which supported the expectancy theory. A multiple regression also carried out as part of the study indicated that the effects of conditioning were completely mediated by expectancies.

This seemed to be convincing evidence in favour of the expectancy theory. However, it is worth noting that some authors (Stewart-Williams & Podd, 2004) argue that conditioning is a mechanism, not a theory, and they claim that conditioning is one method through which expectancies are formed. This theory does not appear to be plausible given the existence of placebo responses which have only been demonstrated with conditioning mechanisms (Benedetti, Pollo, et al., 2003).

2.3.2.1 Expectancy Manipulations

Suggestion (expectancy manipulation) is a feature which while prominent in explanations of hypnotic phenomena, is often neglected in studies of the placebo. This is despite the fact that often the placebo phenomenon is brought about by suggesting to participants that they have received an effective treatment. These suggestions are typically framed as expectancy manipulations; e.g. “This treatment is a potent painkiller which will take effect quickly”. Recently, Kirsch proposed that placebos

could be fruitfully considered in terms of suggestion rather than expectancies (Kirsch, 1999). This viewpoint seems illuminating, as there are large differences in the size of placebo effects depending on the type of expectancy manipulation used.

This line of research began with Kirsch (Kirsch & Weixel, 1988) when he looked at the effects of either telling participants that they would receive coffee, or that they might receive coffee. This clever design mirrors the difference between placebo studies and double blind trials. This experimental study found that when coffee was deceptively administered, there was a much larger effect. Additionally, the physiological parameters measured showed effects in the opposite direction between these two conditions (c.f. Chapter 7).

This finding has been replicated in a clinical setting using analgesics following surgery, showing that deceptive administration produced larger effects than did double-blind administration of the same placebo (Amanzio, Pollo, Maggi, & Benedetti, 2001), though another author failed to replicate this finding using student experimenters (Walach, Schmidt, Dirhold, & Nosch, 2002). In the Amanzio *et al* study, 30 surgical patients were enrolled in either a natural history condition, a double-blind condition or a deceptive administration condition. The outcome variable was the amount of painkillers requested over the recovery period. Each group was significantly different from each other, and the effect size from the ANOVA was $F = 0.92$. Note that this effect size is for the difference between groups, and was computed assuming that the study possessed 80% power⁴.

These research findings argue in favour of suggestion (or alternatively, expectancy manipulation) with regards to placebo being one of the major factors in driving the effect. Additionally, other authors have suggested that suggestion and placebo have much in common, and the lack of linkages between them may be due to lack of clarity in definitions (De Pascalis, Chiaradia, & Carotenuto, 2002).

Furthermore, some research in hypnosis indicates that even the features associated with hypnosis (lack of memory, lack of volition etc) are themselves the result of personal and cultural suggestions (Kirsch, 1999). This does seem to suggest that hypnosis and its effects are just as subject to suggestion as any other interpersonal phenomenon. It may be that placebo effects are merely the result of suggestions (conscious or unconscious) which are given in the domain of health, while hypnosis merely refers to suggestions given in the context of hypnotic treatment or entertainment.

Additionally, there is some meta-analytic evidence that in RCT's at least, expectancies are not always associated with improvement (DiBlasi et al., 2001). This systematic review found that in only ten of the nineteen studies meeting the inclusion

⁴which will have been biased by the filter supplied by the publishing process

criteria were greater expectancies associated with larger placebo effects. The same author also found that many participants in clinical trials did not consider themselves to have expectancies around improvement at all, instead commenting about their hope.

It is worth noting however, that the expectancy theory is a psychological construct mainly of interest to researchers in the field, and even if expectancies explain the entire variance in placebo response, then it would still not be necessary for participants to think about their hopes or fears in terms of expectancy in order for this construct to be useful in describing their response to treatment in clinical or experimental studies.

However, in light of the research reviewed above, and following Stewart-Williams, it would seem most useful to conceptualise suggestion as a means through which particular sets of expectancies are generated, and this approach was taken throughout this thesis.

In conclusion, expectancies are regarded as being an important factor in the response to placebo, and expectancy manipulations are used in many studies to induce placebo responses in participants. However, there is some research that suggests expectancies are not always important (Geers, Weiland, et al., 2005) (c.f. section below), and the measurement of expectancies is currently quite crude. It was this finding (discussed further in Chapter 5) that lead to this thesis focusing attention on the development of a more comprehensive measure.

2.3.3 Motivational Theories

A competing perspective on the placebo has been advanced recently by Michael Hyland (Hyland et al., 2006). Hyland's theory is called motivational concordance, and it regards the behaviours which people engage in and the meanings that they attach to these as primary, rather than the cognitive focus of response expectancy theory.

His research seems to show that depending on how a particular therapy is framed, different variables can predict the placebo response. In the study cited above, spirituality predicted the placebo response to Bach herbal essences, while expectancy was not an independent predictor. In further research (Hyland et al., 2007) he established that this is only the case when flower essences were framed as a spiritual treatment, and not when they were described as motivational tools. These findings argue against the contention of Kirsch that expectancies mediate the effects of placebo on the body directly, or at least suggest that particular forms of expectancies are more effective than others.

He also noted that when a placebo sleep therapy (which involved writing down things

which participants were grateful for) was utilised, gratitude was the best predictor, and again, expectancy added nothing to the results (a finding also noted by Geers (Geers, Helfer, et al., 2005)).

These findings are quite interesting, as they imply that although (explicit) expectancies may provide a useful framework through which to understand placebo effect, they are not a sole explanation.

One issue with the theory of Hyland is that it has not undergone extensive testing, and has never been examined under double-blind conditions, so it is unknown at this point to what extent it will generalise across the various conditions of placebo administration. For example, his results could be due to experimenter effects and demand characteristics. This explanation seems less likely given the results of a randomised study showing that this theory predicted placebo response better than the expectancy theory (Gaitan-Sierra & Hyland, 2011).

Another recent theory regarding placebo effects is that of Geers (Geers, Weiland, et al., 2005). Geers et al notes that motivational approaches to placebos were popular in the past. He suggests that this perspective may prove fruitful for an analysis of placebo effects. His research used priming techniques in order to influence the desires of participants to respond to the treatment.

The major finding of Geers *et al* across a number of studies (Geers, Kosbab, Helfer, Weiland, & Wellman, 2007; Geers, Weiland, et al., 2005) was that placebo effects were significantly greater after participants had been primed with cooperative goals. His research also showed that expectancies had an impact, but again it was not independently significant after motivation was controlled for in a step-wise multiple regression procedure.⁵ Effects of motivation were also demonstrated by Jensen and Karoly, (Jensen & Karoly, 1991). The Jensen *et al* research found that motivation was a predictor of placebo response while expectancies were not, while the Geers *et al* studies found that motivation and expectancies interacted to produce the observed effects.

Some neurological research does suggest that goal directed behaviour may be associated with endogenous dopamine release (Scott et al., 2007) which could provide a plausible mechanism through which goals and motivation help to activate placebo effects. Some other evidence that would support the theory of Geers is that patients suffering from illness typically show much greater placebo responses to experimental pain than do healthy controls (Klinger et al., 2007). The research of Geers *et al* used various different priming manipulations to increase motivation to respond, which also suggests that implicit (or unconscious) motivations may be able to influence the response to placebo.

⁵But such an approach is prone to over-fitting in the absence of an independent test-set

2.4 Moderators of the Placebo Effect

Moving on from theories about the nature of the effect, the next step is to examine factors which can moderate the placebo responses observed in research and clinical practice, covering both experimental data and the results of large meta-analyses.

This section will address potential moderators of the effect. These moderators will be divided into factors inherent to the participant, factors relating to the health care provider, and factors relating to the nature of the placebo and the study design. This review of placebo moderators is focused on individual-difference variables as this was the focus of this research. Note that optimism and its relationship to placebo is covered in Chapter 4.

2.4.1 Certain and Uncertain Expectations

Perhaps the most important feature of trial design affecting the response to placebo is the influence of suggestion/expectancies. In most clinical trials, participants are informed that they will receive either active treatment or placebo. Some authors have suggested (Kirsch & Weixel, 1988) that this process diminishes expectancies related to the treatment efficacy, which in turn reduces the effects (Kleijnen, De Craen, Van Everdingen, & Krol, 1994). The Kirsch study noted above looked at the effects of differing instructions on the results of ingesting placebo caffeine, and showed larger effects when participants were given placebo coffee with suggestions that it was real than when they were told there was a 50% chance they would receive placebo. A replication attempt by Walach *et al* did not confirm this finding, although this study was focused on experimenter effects on placebo, and was not powered to detect this difference.

More recently, Amanzio and colleagues (Amanzio *et al.*, 2001) replicated this finding with patients recovering from thoracic surgery. The principal finding was that those patients who believed they were getting a real medicine required much less analgesia than those who believed that they might receive placebo. Such an effect could account for the differences in effect sizes seen between experimental and clinical studies of placebo. Indeed, the results of Amanzio *et al* suggest that variations in placebo response are responsible for much of the variability in the response to analgesics in general.

A second, related factor may be the use of suggestions in the experimental research. Participants are typically told that they will receive a powerful painkiller before placebo administration, whereas in the clinical trial, no such instructions are given. This finding regarding certain and uncertain expectations was also replicated in a test of a placebo sleep therapy by Geers (Geers, Weiland, *et al.*, 2005).

Linking to the discussion above regarding certain and uncertain expectations, it may also be important to examine a paper by Ploghaus *et al* (Ploghaus, Becerra, Borras, & Borsook, 2003) where the authors argue that certain expectations of aversive events are associated with fear, while uncertain expectations are associated with anxiety. Anxiety is associated with both the nocebo effect and the production of the hormone CCK, which may be why uncertain expectancies appear to lead to lower placebo effects (Colloca, Sigauco, & Benedetti, 2008). These two emotions activate differing parts of the brain, and given the finding that dopamine systems are activated differentially by certain and uncertain expectancies (Scott *et al.*, 2007), this may point towards some important future avenues for research. This new focus on the brain and body correlates of placebo effects has contributed much to the field, as will be seen below in section 2.6.

2.4.2 Treatment Preference

There is some evidence that the benefits accruing from clinical trials may result from the patients expectancies about whether or not they have received the real treatment (Bausell *et al.*, 2005). This study showed no difference between sham and real acupuncture but showed large differences between the outcomes of those who believed they received real treatment versus those who did not. This factor is typically ignored in clinical trials, although prominent commentators have argued that it should be taken more into account (Benedetti, 2007). This finding was later replicated (Linde *et al.*, 2007) where four clinical trials of acupuncture were pooled. This study suggested that although real acupuncture showed similar improvements regardless of expectancies, the minimal acupuncture groups improvement was dependent on their expectancies around acupuncture and perceived treatment assignment. In another trial, belief in receiving real nicotine replacement therapy was an extremely good predictor of successfully stopping smoking (Benedetti, 2008).

These results do suggest that pre-existing expectancies and treatment preferences and beliefs about assignment are an important moderator of the observed placebo responses.

2.4.3 Price of placebo

Another factor which may affect the placebo response is price (Shiv, Carmon, & Ariely, 2005). A Shiv *et al* study utilised an energy drink distributed to college students in an on campus gym under two conditions. In the first, they merely received the drink and were asked to solve a number of puzzles. In the second, they received the same drink, but were told that the price had been discounted (without being given a reason). The participants in the second condition solved significantly less

puzzles than those in the first, suggesting an impact of perceived price on the effectiveness of the energy drink.

This finding has been replicated in placebo analgesia (Waber, Shiv, Carmon, & Ariely, 2008), which is more relevant to this thesis. This finding probably reflects cultural associations of price with value, and one could hypothesise that in other cultures, items perceived as being of greater value would invoke similar effects.

It is worth noting that in neither study were the participants actually required to pay for the drugs, and as such, inferences cannot be made directly regarding the real world effects of these results. Such a study would have much greater external validity and relevance to health care policy makers.

This research also ties into the classic paper by Branthwaite on branded and unbranded pills for the treatment of headaches (Branthwaite & Cooper, 1981). This experimental study, using an extremely large sample, found that branded placebos were more effective than unbranded placebos, suggesting that either advertising or prior learning can affect the effectiveness of two identical preparations.

Again, the findings discussed in this section can be interpreted in terms of expectancies. Price is typically taken as a signal for quality in Western societies, and particular brands of pharmaceuticals and medicines can be associated with relief. That being said, the branding experiments are equally conducive to being explained in terms of conditioning, while the price findings are certainly expectancy driven. Older research did not find any effects of familiarity, which would suggest that the effects of branding have either increased in recent times or that familiarity alone is not enough (Morris & O'Neal, 1974).⁶

2.4.4 Patient/Participant Characteristics

The next kind of moderators of placebo which will be reviewed are psychological characteristics of the individuals under study. Both state and trait variables may be involved here, though most of the research has focused on traits, as they tend to be easier to measure.

2.4.4.1 Somatic Focus

Somatic focus, or the focus on internal bodily sensations (the proprioceptive sense) appears to have an impact on the response to placebo, though this finding has only been demonstrated in a small number of studies.

⁶Or a lack of statistical power

This finding arises from the work of Geers *et al* (Geers, Helfer, Weiland, & Kosbab, 2006) on somatic focus and its effect on the placebo response. In summary, this experimental study asked half the participants to attend to their somatic sensations following placebo administration, and gave the other half no such instructions. A similar finding was made by Rainville *et al* with regard to hypnotic suggestions (Price *et al.*, 2008).

The participants who focused on their bodily sensations showed an increased placebo effect, which is an interesting finding for many reasons. Firstly, it suggests that the effectiveness of a treatment can be increased by asking participants to pay attention, and this notion of somatic focus is one of the ways in which mindfulness is typically conceptualised.

Secondly, it links in with an explanation given for differences in placebo response across treatments following a meta-analytic review (Meissner, Distel, & Mitzdorf, 2007) where the authors provide evidence that placebo effects are not common where the outcome measure is a hormone level, while they are common where the outcome measure is a peripheral disease parameter. They suggest that this occurs because nervous system feedback loops are available for the second kind of outcome, but not for the first. This finding is discussed further in Chapter 3, particularly Section 3.2.3.

2.4.5 Provider Factors

Another factor which is believed to be of importance in placebo effects is the patient-provider relationship. The classic study in this field was performed by Thomas (Thomas, 1994), in which patients suffering from unclear symptoms were given either a positive or negative consultation. The results of this study showed that 2/3rds of the patients given a positive consultation improved, while only 1/3rd of the patients given a neutral consultation had. This study has been called into question by more recent research (Knipschild & Arntz, 2005), but there is some meta-analytic evidence that provider effects may account for a significant portion of the response to placebo (DiBlasi *et al.*, 2001).

In conclusion, the placebo response is a complex phenomenon and can be impacted by internal participant factors, features of the patient-provider relationship, features of the treatment itself, and also features of the setting in which the treatment is administered. Very few studies control all of these factors, and this may contribute to some of the confusion and controversy surrounding the construct.

The above review of the research relating to features associated with response to placebo has established the following.

- Placebo effects appear to be related to expectancies and through these,

optimism (c.f. Chapter 4);

- The perceived value of a placebo (as expressed through price and branding) effects the response (potentially through expectancies);
- The relationship between patient and provider appears to drive some of the observed placebo effects (also potentially through expectancies);
- Treatment preference and beliefs about assignment to treatment appear to be as important as is actual treatment assignment (c.f. Section 5.1.2).

2.5 Effect Sizes for Placebo Studies

The placebo is only amenable to prediction if it can be induced in a given sample. One of the critical requirements for inducing placebo effects is therefore to have some estimate of the size of the effects so that appropriately powered studies can be designed.

One well known meta-analysis suggested that the benefits of placebo were negligible in most areas (Hrobjartsson & Gotzsche, 2001), with the exception of pain trials. While this meta-analysis has been critiqued for its ignorance of psychological studies of placebo (Evans, 2003; Stewart-Williams, 2004) and for the combining of placebo effects across 200 plus treatments (Wickramasekera, 2001), there is no denying its large effects on the field.

One of the major reasons for the popularity of pain studies in placebo research is probably the large effect sizes, as measured by Cohen's d . While effects in some areas range from about $d = 0.15 - 0.25$, the effect sizes in pain studies tend to be much larger, ranging from $d = 0.45 - 0.95$ (Vase, Riley, & Price, 2002). Given that the d measure expresses effect sizes in terms of standard deviations, an effect of between a half and one standard deviation is quite respectable, and allows for smaller studies to examine effects of interest. However, it has been argued that these effect sizes are illusory, and result from lack of blinding, inadequate controls and poor randomisation (Hrobjartsson & Gotzsche, 2003; Kienle & Kiene, 1997).

These two viewpoints can be reconciled, at least in the opinion of Vase *et al* (Vase *et al.*, 2002). In this meta-analysis, Vase and colleagues looked at the sample of placebo pain trials in both the clinical and the experimental areas. What they found was that effect sizes tended to be small when the placebo was used in a clinical trial, and much larger in experimental studies of the placebo effect. This analysis was disputed by Hrobjartsson and Goetzche (Hrobjartsson & Gotzsche, 2003) who noted problems with the methods of analysis chosen by Vase *et al*. Even using the more conservative estimates of Hrobjartsson and Goetzche, the effect size from experimental research

($d=0.5$) is still twice as large as those observed in clinical trials. There are a number of factors which differ in these two contexts which could be responsible for these observed differences.

Sauro *et al*, in a review of endogenous opioids and the placebo effect did find a significant difference between the effect sizes in post-operative and experimental pain, with post-operative pain showing an average effect size of $d = 0.65$, 95 $CI(0.37 - 0.87)$ while the experimental studies showed an average effect size of from $d = 0.53$, 95 $CI(0.02 - 1.04)$ for shock induced pain, to $d = 1.23$, 95 $CI(1.00 - 1.46)$ for ischemic pain (Sauro & Greenberg, 2005), suggesting that ischemic pain is the best way to invoke a substantial placebo effect.

If there is a gap in the literature, it probably results from a paucity of meta-analytic studies on non clinical trials of placebo - with some notable exceptions (Wampol, Imel, & Minami, 2007; Vase et al., 2002), but this is a matter that could be easily addressed by future research. It is however, a matter worth addressing, as what little studies we do have indicate that there are a number of major differences between the data revealed by each of these methodologies.

2.6 Physiology of the Placebo Effect

The placebo effect is an interesting phenomenon in that it straddles the boundaries of psychological and physical. This section will examine the research demonstrating the effects of placebo on a neurological level, and then examine other physiological impacts and correlates of placebo administration. While there are a large number of recent studies examining the fMRI correlates of placebo response in individuals (Benedetti, Mayberg, Wager, Stohler, & Zubieta, 2005), these were not essential to this research, and as such are not reported here.

2.6.1 Opioids and Placebo

The biochemical history of placebo begins with Levine (Levine, Gordon, & Fields, 1978) and the demonstration that naloxone blocks many placebo pain responses. This experimental finding suggested that placebo pain relief is mediated by the endogenous opioid system.

This finding has been qualified by research over the past thirty years, suggesting that both opioid and non-opioid systems can be involved in the placebo pain response depending on the the method of inducing placebo responses and the biological system involved (Amanzio et al., 2001; Benedetti, Pollo, et al., 2003). The lasting contribution of this research is that it paved the way for the placebo to come in from the fringes of medical science.

In this area, the work of Benedetti and his colleagues has been instrumental in unveiling the biochemical pathways through which placebos exert their effects, and much of this work is summarised in his book. It appears that both the opioid and dopaminergic systems are involved in the placebo effect. Benedetti and colleagues have demonstrated that respiratory depression can be induced by placebo administration (Benedetti, Amanzio, Baldi, Casadio, & Maggi, 1999).

A further discovery with regard to placebo analgesia is that it can be directed at specific sites in the body (Benedetti, Arduino, & Amanzio, 1999). This study induced expectancies of placebo responses at either the right or the left hand, and demonstrated the expected placebo effects. These effects were completely antagonised by naloxone, which suggests that they were mediated by the endogenous opioid system.

This finding is interesting as it suggests that the opioid systems can be activated at specific parts of the body, and not just globally as some former theorists have claimed. A more recent finding (Watson, El-Deredy, Bentley, Vogt, & Jones, 2006) found that perhaps 50% of participants in a placebo analgesia study generalised a placebo response across both arms, even though cream was only applied to one arm for each person. This study would suggest that the placebo analgesia phenomenon is quite malleable and subject to individual interpretation (i.e. moderated by “descending” pathways from the brain (Goffaux, Redmond, Rainville, & Marchand, 2007)).

Further research on the blockade of opioid receptors by naloxone has established that proglumide can be used to increase the size of placebo analgesic effects (Benedetti, Amanzio, & Maggi, 1995). Additionally, CCK, a chemical which tends to produce anxiety in human participants, has been shown to increase the size of the nocebo effect (Benedetti, 1996).

A recent meta-analytic review (Sauro & Greenberg, 2005) seems to argue that placebo effects in pain are quite large ($d=.89$) and that naloxone is quite effective in reducing them ($d=.55$), pointing towards an interpretation of placebo effects in pain being substantially mediated by endogenous opioids.

One study which looked at patients suffering from IBS found that naloxone did not reduce the size of placebo effects, which would suggest that these were not opioid mediated (Vase et al., 2005). It remains to be determined why placebo effects in IBS are not opioid mediated, and understanding this may give some insight into the phenomenon.

2.6.1.1 Placebo and Skin Conductance

Skin conductance is another measure sometimes used in placebo experiments. Some authors have reported no difference between groups on this measure (Flaten et al.,

1999), but these authors also examined differences between groups using an ANOVA method. One extremely interesting study claimed that pain ratings could be derived from the measurement of skin conductance, and that active drugs changed the response patterns, while placebo administration did not (Fujita, Fujii, Nakamura, Miyauchi, & Takagi, 2000) (and c.f. Chapter 7).

The Crum et al exercise study discussed earlier also showed physiological effects of their exercise “placebo”, arguing that suggestion and expectancy effects can have sizeable impacts on physiological outcomes (Crum & Langer, 2007).

Placebos have been noted to affect hormone levels also, but these placebo responses appear to be completely insensitive to expectancies, and can only be induced by conditioning (Benedetti, Pollo, et al., 2003).

These kinds of physiological markers of response to placebo are extremely useful as they can be used to determine if a physiological placebo effect is occurring, or if the change in self rated pain is driven by more cognitive re-appraisals of the situation. If more cognitively driven, it would be expected that these changes would lag the changes in self reported pain, whereas if the placebo were mediated locally then it would be expected that the physiological changes would occur in advance of a reported drop in self rated pain. One cautionary point is that biological mediators and correlates of placebo are not the cause of placebo effects, rather they are examples of them, and while the study of physiological correlates of placebo can help to understand how they are mediated, it will not remove the need for social and psychological explanations for how they occur (Stewart-Williams, 2004).

2.7 Review of the Placebo Effect

In this section, the major themes which have emerged from the literature review of the placebo effect are recapped. There are a number of major points to bear in mind.

Firstly, the placebo effect is a difficult phenomenon to define precisely, and there is wide variability in what is considered to constitute a placebo effect. One of the clearer conceptualisations of the effect is that proposed by Goetzche (Goetzche, 1995), where it is argued that placebo effects should be broken down into three parts: those attributable to the patient-provider interaction, those attributable to the administration of the medicine, and those attributable to the context in which the treatment was delivered, and these three factors have been taken into account in the definition given in Section 2.2.4.

The next major theme to emerge from this review is that there is one theoretical perspective shared by most researchers in the field, that of response expectancies (Kirsch, 1997, 1985). However, this theory, while still core to the

conceptualisation of the effect, has been contextualised by a number of demonstrations that expectancies are not always the best predictor of placebo (Hyland et al., 2006; Geers, Weiland, et al., 2005). This theory also suffers from the lack of clear terminology and instruments with which to measure expectancies, a deficiency which this research hopes to remedy.

The major factors relating to context appear to be the setting in which placebo is administered (clinical trial or experimental), along with the participants beliefs about treatment assignment. Another major contextual factor appears to be the rationale given for the placebo's effects. In addition, provider factors appear to include the level of suggestion given (certain versus uncertain) and the charisma and authority of the provider. Patient characteristics which have been shown to effect the response to placebo include dispositional optimism, somatic focus, and in some situations, gratitude and spirituality.

The meta-analytic evidence, though conflicting, appears to indicate that placebo effects occur when the outcome variable is under the control of the central nervous system, and do not occur nearly as much in the endocrine system. However, this finding is based on clinical trial data and contradicts the successful conditioning of humans to respond to endocrine placebos (Benedetti, Pollo, et al., 2003). The size of placebo effects is also a matter of some dispute, and appears to significantly differ as a function of study design, as noted above in Section 2.5.

In conclusion, the placebo effect is a complex phenomenon which appears to provide a link between the psychological and physiological experience of the world, and which is associated with some psychological and physiological variables.

This review has established the following:

- The placebo effect seems to occur in many situations, in both experimental and clinical trials;
- Expectancies and optimism have been connected to the placebo response;
- As yet, there is no multi-item scale for assessing the treatment-related expectancies associated with placebo response;
- The placebo effect is typically associated with a lack of awareness relating to the treatments veracity.

This thesis contributes a measure of treatment expectancies, and an investigation of both measuring the placebo effect with an implicit measure, and the relative usefulness of both expectancies and optimism to the prediction of this construct.

2.8 The Implicit Association Test

A brief introduction to implicit measures in general, and the IAT specifically was given in Chapter 1. Given that the focus of this work was on the development of an IAT to measure treatment credibility, this section focuses on the psychometric features of the instrument, its predictive validity and the settings in which it has been used.

2.9 Psychometric Analysis of the IAT

The method has become very popular, and has been applied to many areas of social psychology, such as attitudes towards fatness (Ahern, Bennett, & Hetherington, 2008), towards disability (Pruett & Chan, 2006) and towards smoking (Kahler, Daughters, Leventhal, Gwaltney, & Palfai, 2007). However, regardless of popularity, a measure is only useful to the extent that it possesses the following three properties:

1. Validity - it measures what it purports to measure;
2. Reliability - the measurements are consistent across time;
3. Predictive power - unless the measure predicts some outcome of interest, it is essentially useless.

In this section, the validity and reliability, along with the predictive capability of the IAT will be examined to determine if this measure is likely to prove useful in this thesis.

2.9.1 Convergent Validity

One of the major elements used to define and assess the validity of a measure is that of convergent validity, which assesses the strength and directionality of relationships between the measure and other constructs which are theoretically related to the measure or construct. In this section, the relationships of the IAT to other explicit measures and implicit measures will be examined.

2.9.1.1 Relationships with Explicit Measures

IAT scores are typically weakly correlated with explicit measures of similar attitudes. These correlations average ($\bar{r} = 0.39$) (Nosek, Greenwald, & Banaji, 2005), and so one could be justified as regarding the two as distinct constructs (Nosek & Smyth, 2007) and this is the approach taken by many of the originators and early workers in the field (Greenwald & Farnham, 2000; Nosek & Smyth, 2007).

Additionally, they tend to reveal stronger associations than explicit measures when the topic is politicised or controversial (Greenwald et al., 2009), but tend to provide less predictive power in less controversial situations, such as self-reported soda choice (Karpinski & Steinman, 2006).

It has been proposed that implicit measures should be more resistant to change than explicit measures (Greenwald & Banaji, 1995; Greenwald et al., 1998). Recent research has shown this not to be the case (Meagher & Aidman, 2004; Gschwendner, Hofmann, & Schmitt, 2008). In these experiments, IAT scores were shown to change much more than explicit measures in response to experimental manipulations. This may suggest that the measure captures state rather than trait variance, a position which will be discussed further throughout this thesis.

Dasgupta (Dasgupta & Greenwald, 2001) determined that implicit evaluations could be significantly changed by displaying either positive or negative exemplars of the construct under study. Dasgupta *et al* suggested that this may mean that implicit associations are more effected by shallow processing techniques, which is a theory entirely consistent with the notion that they capture state variance.

One major requirement for convergent validity is that the measure should correlate with other measures known to tap similar constructs. This requirement has been fulfilled for the IAT. One study showed that mindfulness (as measured by the MAAS) was a mediator of the relationship between explicit and implicit attitude scores (in the particular study concerned, attitudes towards autonomy) (Levesque & Brown, 2007). Additionally, private self consciousness, an analogous construct, was found to also act as a moderator of the relationship between implicit and explicit attitudes (Gschwendner, Hofmann, & Schmitt, 2006). Introspection was negatively related to the correlation between implicit and explicit measures (Hofmann, Gawronski, Gschwendner, Le, & Schmitt, 2005), while spontaneity was positively associated with this correlation.

Another feature which moderates the relationship between implicit and explicit attitudes is that of attitude importance. As attitudes increase in importance, the correlation between explicit and implicit attitudes increases, suggesting that implicit attitudes tap cognitions that can be, but are not necessarily, conscious (Karpinski, Steinman, & Hilton, 2005).

It appears that IAT's may predict spontaneous behaviour better than explicit measures do (Asendorpf et al., 2002; Richetin, Perugini, Prestwich, & O'Gorman, 2007; Perugini, 2005). This has been demonstrated in a number of domains, and appears to suggest a so-called double-dissociation model, whereby (in a food example) IAT measures predict spontaneous food choice while explicit measures predict food diary measures (Richetin et al., 2007).

The double dissociation model has been tested using Structural Equation Modelling (Nosek & Smyth, 2007; Perugini, 2005) and appears to provide a better fit than a pure explicit or pure implicit measure.

The form of the explicit measure also appears to affect the relationship between explicit and implicit measures. It appears from a meta-analysis that relative explicit measures tend to have a far greater correlation with implicit measures than do explicit ones (Hofmann et al., 2005). This makes sense given that the IAT is a relative measure. For this reason, the measure developed in Chapter 5 is not a relative measure, in order to attempt to maximise the incremental variance explained between the explicit and implicit measures.

2.9.2 Predictive Validity of the IAT

The major proposed advantage for the use of implicit measures is that they would act as better predictors of behaviour, or allow for more insight into hidden cognitions that were associated with behaviours yet either not accessible or not reported by participants (Greenwald et al., 1998). This section examines the extent to which these hopes have been fulfilled. While the IAT does not appear to be a better predictor of behaviour overall, it does possess some ability to predict behaviours which are typically hard to predict using self report measures (Greenwald et al., 2009).

The classic demonstration of the difference in prediction between implicit and explicit measures relates to (Asendorpf et al., 2002) who investigated the attitude of shyness. In this study, spontaneous shy behaviour was predicted by implicit associations, while controlled shy behaviour was predicted by explicit attitudes. This pattern has become known as double dissociation, and has been observed in a number of studies (Perugini, 2005), and also supported by a meta-analysis (Hofmann et al., 2005).

A recent review of implicit measures of self esteem suggests that implicit and explicit self esteem are entirely distinct constructs (Rudolph, Schroder-Abe, Schutz, Gregg, & Sedikides, 2008). Implicit self esteem has been shown to predict response to success or failure (Greenwald & Farnham, 2000). The research on implicit and explicit self esteem seems to indicate that individuals can be classified as having particular types of self esteem based on their relative levels of implicit and explicit self esteem, where participants who have high levels of both implicit and explicit self esteem are classified as having genuine self esteem (Meagher & Aidman, 2004).

These participants tended to be more resilient and suffer less negative outcomes following a false feedback manipulation designed to reduce self esteem. Although the effect size for the explicit measures was much higher in this study, the implicit measures (the Self Apperception Test and the IAT) appeared to be more sensitive to the emotional tone of the feedback.

In the domain of personality, implicit measures of all Big 5 traits have been correlated with spontaneous behaviour which reflected these traits (Steffens & Konig, 2006). The Steffens *et al* study predicted behaviour based on the Big 5 personality traits. There was no effect for extraversion, but the experimental manipulation may have been poorly designed. This findings lends more credence to the double dissociation theory (Asendorpf et al., 2002) which has been shown to apply in a wide variety of tasks since then (Perugini, 2005; Conner & Barrett, 2005).

One paper (Boldero, Rawlings, & Haslam, 2007) which used the Go/No Go Association Test (GNAT) showed that implicit Extraversion and Neuroticism were able to predict reaction time in the experiment. More generally, the implicit attitudes were able to predict scores on the explicit attitude measure.

However, the Big 5 traits predicted were different from the findings of other researchers, suggesting that there may be some method variance involved. This method variance has also caused problems for the IAT (Mierke & Klauer, 2003; A. G. Greenwald, Nosek, & Banaji, 2003). Again, this highlights the issue that these new implicit measures may have confounding factors within them which can only be highlighted by further detailed research.

Gender IAT scores have been shown to correlate with ratings of prejudiced behaviour in a simulated job interview task (Rudman & Glick, 1999). Additionally, poor interactions with an opposite race experimenter have been associated with high scores on the Race IAT (McConnell & Leibold, 2001). There is some evidence that IAT measured preference for males over females can result in prejudicial behaviour against females in a simulated interview setting (Greenwald & Farnham, 2000; Heider & Skowronski, 2007). The magnitude of the IAT scores was correlated with the observer-reported prejudicial behaviour scores.

Some have critiqued these findings as both IAT and behavioural assessments were carried out in the same session, and this may falsely inflate the attitude-behaviour correlation. More recent research has demonstrated that even when separated by a week, attitude assessments using the IAT are significant predictors of verbal and non verbal friendliness with a compatriot of opposite race (Heider & Skowronski, 2007).

A further study on explicit moderators of the predictive validity of the IAT showed that participants scores on the Self Reported Habit Index moderated the predictive validity of the IAT (M. T. Conner, Perugini, O’Gorman, Ayres, & Prestwich, 2007). This finding, along with the relationship between mindfulness, introspection and other similar constructs and IAT scores, provides evidence that IAT scores are differentially predictive based on moderator variables.

The predictive validity of the IAT may also vary as a function of domain. An Italian study (Arcuri, Castelli, Galdi, Zogmaister, & Amadori, 2008) showed that the IAT

was able to predict the future voting behaviour of people based on the results of an IAT measuring attitudes towards left and right wing candidates. This finding is quite impressive, given the difficulties political scientists normally find in predicting the behaviour of undecided voters. However, this should be qualified with the fact that it was self reported voting behaviour was measured, as opposed to actual voting patterns. Although the Arcuri findings are interesting, other research on the relationship between IAT scores and voting behaviour showed that IAT scores had no incremental validity over explicit measures in prediction of vote choice (Karpinski *et al.*, 2005). The Karpinski *et al.* study did not focus on undecided voters, however, which may be responsible for the differences in the results.

Other research has indicated that IAT scores are more predictive when participants are under strain or otherwise tasked for resources (Hofmann, Gschwendner, Castelli, & Schmitt, 2008). These findings suggest that IAT scores typically have quite large amounts of state variance if they are malleable in terms of experimental manipulations such as this. An alternative theory is that the IAT responses can be inhibited by awareness, which caused the observed differences in the studies reported above.

A study which demonstrated a link between chronic pain and implicit measures of pain and identification with pain was reported in Grumm *et al.* (Grumm, Erbe, von Collani, & Nestler, 2008). This study followed a group of chronic pain patients as they took part in a mindfulness-based intervention for pain reduction. The use of pre-post design allowed for the authors to assess that the intervention caused a significant drop both in self reported pain and medication use, and a reduction in self-pain associations as measured by an IAT. This study is one of the first to map an IAT to a health-related construct, and to actually take behavioural measures into account.

To conclude, there is substantial evidence for the predictive power of the IAT across a number of disparate domains, but there is much more work to be done given that predictive power seems to be quite malleable across experimental studies and there is some conflicting evidence, especially in the area of political attitudes.

2.9.3 Reliability of the IAT

The reliability of the IAT has been a major issue throughout its existence. Test-retest reliability seems to average around ($r = .49$) which, while permissible in a psychometric instrument for theoretical purposes, is far too low for making clinical or legal judgements (Greenwald & Farnham, 2000; Blanton & Jaccard, 2006). It is worth noting however, that the test-retest reliabilities do not drop much farther than this, even over periods as long as one year (Egloff, Schwerdtfeger, & Schmukle, 2005). Split-half reliabilities are typically higher, averaging around ($r = 0.80$),

This may suggest that the IAT response is composed of both state and trait portions,

and that the trait portion of the measure is relatively invariant across temporal distance. Some authors have argued that this low test-retest reliability is due to both error variance and person by situation interactions (Gschwendner et al., 2008). Manipulation of accessibility of the constructs measured in the IAT has been shown to improve the temporal stability of the IAT scores, suggesting that particular situations may make the constructs more or less likely to be expressed (Gschwendner et al., 2008).

2.10 Moderators of the IAT

2.10.1 Contextual Moderators of IAT Effects

Some contextual factors have been noted to affect IAT measures, even though these were some of the problems which the measure were designed to avoid. A study asking participants to complete the IAT under conditions when they believed that the experimenter would or would not know their scores (the so-called bogus pipeline) (Boysen, Vogel, & Madon, 2006) showed a notable diminution in IAT effects in a measure of attitudes towards homosexuals when participants believed that the experimenter would know their scores. Another study examining the attitudes of Italian students to Turkish immigrants replicated this finding, with IAT scores being reduced when in the presence of others (Castelli & Tomelleri, 2008).

2.10.2 Cognitive Moderators of IAT effects

Another factor which appears to moderate the observed IAT effects is that of memory resources. A study looking at attitudes towards Blacks and Turks (Hofmann et al., 2008) found that the IAT acted as a far better predictor of behaviour when participants had been asked to remember a list of words than when they were untaxed. This finding was replicated by the same authors using a different sample and IAT, suggesting that it is quite robust. This suggests that the attitudes measured by the IAT are the result of more automatic processes, and will have predictive power to the extent that the matter involved is not the subject of deep processing (Kahneman et al., 2002).

The Hofmann (2008) study described above involved interaction with an experimenter of the out-group measured in the IAT, and the second study separated the IAT from the behaviour assessment by one week, so these results are both good measures of behaviour and unaffected by issues of attitude-behaviour consistency .

Some research (Dasgupta & Greenwald, 2001) suggests that showing exemplars of groups typically the subject of negative associations on the IAT measures (such as

Black Americans or Females) can reduce the size of these associations. However, as described below, difficulty in recalling such exemplars can lead to larger IAT effects.

It also appears that IAT effects are influenced by ease of retrieval mechanisms (Kahneman et al., 2002). An extremely well conducted study examined a number of implicit measures and the mechanisms through which they are influenced by context (Gawronski & Bodenhausen, 2005).

Gawronski *et al* class the IAT as a response compatibility measure, and argue that these measures are affected by ease of retrieval from memory. In support of this, participants who generally liked African Americans showed higher levels of implicit preference against this group when asked to generate a high number of either liked or disliked African Americans. Conversely, participants who generally disliked African Americans showed lower levels of prejudice when they generated a lower number of exemplars. The authors explain this effect in terms of ease of retrieval. The subjective difficulty of generating the exemplars seems to alter the attitudes which the participants report (Kahneman et al., 2002).

Another study (Levesque & Brown, 2007) looked at the effects of mindfulness on expression of implicit attitudes and argued that high mindfulness can stop implicit attitudes from being expressed and over time, causes them to become more in tune with self reported attitudes. This study used an experience sampling methodology and examined attitudes towards autonomy and heteronomy. The findings suggested that participants high in mindfulness seemed to show higher levels of autonomy in general, and that mindfulness could act as a protective factor against the expression of unwanted implicit attitudes. This finding has also been supported by other recent research (Gschwendner et al., 2006), when they noted that Private Self Consciousness seemed to correlate with the expression of implicit attitudes towards Germans and Turks. In addition, the self reported habit scale, a measure which has been found to be negatively correlated with mindfulness, was found to be associated with stronger implicit attitudes and less congruence between explicit and implicit attitudes (M. T. Conner et al., 2007).

2.11 Conclusions

In summation, the Implicit Association Test is a new and useful measure. It has been embroiled in a number of spirited debates since its publication, and much progress has been made in the understanding of the measure as a result of this.

The measure can reliably differentiate between groups possessing different attitudes, and seems to produce results which while correlating somewhat with explicit attitudes, appear to reflect a different underlying process (Nosek & Smyth, 2007).

It can predict behaviour quite well in certain situations, especially in matters of importance to the participants and when they are working with limited cognitive resources. It appears to predict spontaneous behaviour much better than explicit measures, which is certainly of interest. This prediction of spontaneous behaviour, coupled with low test-retest reliabilities, would seem to suggest that there are major state components to the measure.

However, the measure still has its problems. There is still no generally accepted theoretical rationale for its effects, it can be contaminated by such issues as task switching costs (Klauer & Mierke, 2005), processing speed (Blanton et al., 2006) and the context in which it is administered (Boysen et al., 2006).

The D algorithm (described in Chapter 7) seems to do a good job of controlling for task-switching costs. Secondly, valence and familiarity need to be controlled for in order to avoid contaminations with general processing speed and salience asymmetries. Thirdly, the measure does not have a theoretical foundation, and this is something that needs to be addressed.

One can argue that if it works then it should be utilised, but a cogent theoretical account of the IAT would allow for more detail and precision in experimental design than is the case at present, and for this reason such a development is essential if the IAT is to become a permanent part of research in experimental psychology.

2.12 Placebo: Implicit and Explicit Expectancies

So, having reviewed the placebo effect and the implicit association test above, the main experimental hypothesis of this thesis can now be examined: that the placebo effect can be predicted by a combination of self reported and implicitly measured (explicit and implicit) expectancies. Below, the reasons for believing this shall be set out, and in the next section research evidence shall be presented which supports this original hypothesis.

Individual level predictors of the placebo response have proved elusive in experimental work for over 50 years. Many psychological variables have been tested to see if they can predict the response, but almost all have failed to replicate (Shapiro & Shapiro, 1997). Indeed, some argue that placebo responders do not even exist (Kaptchuk, Kelley, Deykin, et al., 2008). This thesis contends that the lack of results in the prediction of the placebo effect is an artefact of the methods used rather than the unpredictability of the effect. One of the contentions of this thesis is that the Implicit Association Test is a useful method with which to predict the response.

2.12.1 The research evidence for the combination of the two measures

Research on placebo focusing on either conditioning or expectancy explanations lead to a number of findings which suggest that implicit measures may be a way in which placebo responses can be predicted. In this section, the expectancy versus conditioning debate shall be reviewed insofar as it relates to the main point, which is to examine whether or not the use of implicit measures of expectancies will allow us to predict the placebo response in individuals with greater accuracy. Then some other lines of evidence which support this case will be reviewed, and finally the evidence for this hypothesis shall be reviewed.

The general format developed by Voudouris and still in use today in the research of Benedetti (Benedetti, Amanzio, Vighetti, & Asteggiano, 2006) and others, consists of three stages.

Firstly, the participant has their pain thresholds calibrated and is given a number of blocks of painful stimuli. Secondly, the participants are either given the same stimuli again following the application of placebo cream or the pain is reduced for the second stage after the application of the cream. The second group are then said to be conditioned by this stage. In the third block, pain is increased again for the conditioned participants, and they typically show a much larger placebo response than those who are merely given verbal suggestions of analgesia.

These results were believed by Voudouris to argue in favour of a conditioned approach to the placebo effect. However, an experiment by Montgomery and Kirsch (Montgomery & Kirsch, 1997) demonstrated that these conditioned effects resulted solely from expectancies, and after a regression in which expectancies were partialled out, there were no significant effects arising from conditioning. This would seem to argue that expectancies are the prime method through which conditioning has an impact, and indeed this is the position of Kirsch.

Kirsch's (Kirsch, 1985, 1997) theory of response expectancies specifies that these are the expectation of a non-volitional response, and he argues that they play a role in placebos and hypnosis. This theory relies upon the measurement of self report expectancies to determine this, and this seems to be somewhat incompatible to what happened in the Montgomery and Kirsch research described in the paragraph above. In the experiment that confirmed the effects of expectancies, there were two groups who received lowered stimuli to condition them. One of these groups was informed of this pairing, and the other was not.

Contrary to the predictions of the conditioned response model, those in the informed pairing group did not show an enhanced placebo effect, while those in the uninformed pairing group did. This seems to indicate that the effects of the conditioning

procedure were inhibited by awareness. In other words, what occurred here could be described as an example of implicit learning (learning without conscious awareness).

This conclusion is further reinforced by the work of Shiv and Carmon (Shiv et al., 2005) in the placebo effects of energy drinks given to participants at a lowered price. This research used an outcome measure of the number of problems solved in a specified period, and it demonstrated that those who believed that the price had been discounted solved less puzzles than those who received the drink at the normal price. The interesting feature of this research (for the purposes of this thesis) is that when participants attention was drawn to the discounted price, the difference between the groups was reduced. This would seem to argue in favour of an interpretation involving learning without conscious awareness.

From the research into the IAT, it appears that the measure captures state variance, and given the non-replicability of the placebo response across participants and conditions (Whalley, Hyland, & Kirsch, 2008; Shapiro & Shapiro, 1997), this would also seem to be a factor in the placebo response. Secondly, the research on the IAT suggests that it is likely to be more predictive in contexts where participants are depleted of resources - such as when they are sick, or in pain, which are common conditions in which placebo effects are observed.

So, reviewing these research findings, we can note that in some cases, the placebo effect seems to emerge without conscious awareness. The Shiv *et al* (2005) study noted above showed that when participants awareness was drawn to the discounted price, the effects disappeared. A similar phenomenon occurred in the Montgomery and Kirsch (1997) (Montgomery & Kirsch, 1997) study. These research findings suggest that the placebo response is at least partially determined by factors outside conscious awareness, and as such, it would make more sense to use implicit measures, of which the IAT is the most prominent.

One line of evidence which supports this thesis is the finding that the IAT is better at predicting spontaneous behaviour than explicit measures (Conner & Barrett, 2005; Hofmann et al., 2005). The IAT outperforms explicit measures in some domains (Greenwald et al., 2009) and these domains tend to be where there is little conscious deliberation or reflection upon the matter concerned. The placebo effect is the example par excellence of a un-deliberated and spontaneous type of phenomenon, as no one chooses to have such a response and it seems dependent upon the lack of awareness of a participant that the treatment which they are receiving is not an active one. Thus, we can take this as supporting evidence that implicit measures should predict the placebo response more effectively.

In conclusion, the evidence reviewed in this chapter seems to show that the expectancies underlying the placebo effect can be measured by means of the Implicit Association Test for the following reasons.

- Firstly, some placebo effects seem to require a lack of conscious awareness in order to occur (Shiv et al., 2005; Geers, Weiland, et al., 2005).
- Secondly, the placebo effect seems to be best modelled as a spontaneous phenomenon and the IAT has been shown to predict these kinds of behaviours better (Asendorpf et al., 2002; Richetin et al., 2007).

For these reasons, the development of an IAT which can measure these expectancies is a worthy contribution to human knowledge.

Chapter 3

Theory & Methodology

3.1 Introduction

This research assessed the differential contributions of implicit and explicit expectancy measures of treatment credibility and optimism to the prediction of the placebo response. In this section, the theory underlying this primary research question will be explicated, and followed by a description of the methods used to answer the primary research question.

This chapter will consist of the following sections, representing the core parts of the thesis.

1. The theoretical model of the thesis will be explained;
2. The methods used for the development and validation of the scales used in the research will be described.

The development of the implicit measures is described in Chapter 6.

3.2 Introduction to the theory

Theories form an indispensable part of science. They represent an attempt to generalise beyond particular forms of evidence and data and to derive some kinds of overall principles which lie behind the observed events. The major theories behind the placebo and implicit measures were reviewed in Chapter 2, and in this chapter provides an attempt to synthesise all of this information into a coherent whole. The major building blocks of this theory are as follows:

1. The research base into implicit measures point towards there being at least two systems of attitude assessment and evaluation of stimuli in the human mind;
2. The placebo effect is typically conceptualised as a conscious phenomenon, in spite of the experimental evidence;
3. There appear to be feedback loops between bodily sensations and conscious perception i.e. embodied cognition;
4. These feedback loops are evidence against an additive model of drug-placebo interactions;
5. Despite the centrality of such conceptions to the placebo effect, no theory has as yet incorporated these findings into their theories.

This section will briefly review the evidence in favour of the above propositions, then will elucidate how these could be combined into our theories of placebo and implicit measures, and will provide some testable hypotheses regarding the theory, in the spirit of Popperian falsification.

3.2.1 Implicit Measures and Dual Process Models of Mind

The notion of dual process models of mind is an old one, dating back within psychology to at least the time of Freud, and possibly before. However, the modern conception of dual process models is much more recent, and developed as a result of work with implicit measures and through the findings of cognitive psychology (Kahneman et al., 2002; Greenwald & Banaji, 1995; Gigerenzer & Gaissmaier, 2011; Klauer, Voss, Schmitz, & Teige-Mocigemba, 2007). Essentially, the modern theory suggests that there are two prevalent systems of reasoning inherent to humans, a slow, rational, conscious system, and a fast, frugal and implicit system (Kahneman et al., 2002).

These two systems activate under different conditions and seem to perform different functions. One of the major hypotheses emerging from this theory is that under conditions of attentional strain, the implicit system takes over. This is, as has been seen in the previous chapter, borne out by much of the research into implicit attitudes. Asendorpf (Asendorpf et al., 2002) demonstrated what came to be called the double-dissociation effect, where implicit measures (of shyness, in this case) were more predictive of performance on a spontaneous speaking task (the Trier Social Stress Test), while explicit measures were more predictive of considered, deliberate behaviour. These findings have been further replicated by other authors (Richetin et al., 2007). Thus, results of IAT research and that into other implicit measures can be taken as pointers towards the operation of this system. The major point to take from this section is that this system appears to exist, and yet has only been touched upon in one or two articles over the past decade (Geers, Helfer, et al., 2005).

3.2.2 Conscious Conceptions of the Placebo

The dominant model within placebo research at present is the response expectancy theory of Kirsch (Kirsch, 1985, 1997). This theory conceptualises the placebo as resulting from response expectancies, which are defined as “the conscious expectation of a non-volitional response”. This theory has had some success, displacing the then prevalent theory of conditioned placebo responses (Voudouris et al., 1985; Voudouris, Peck, & Coleman, 1989). That being said, the very definition of placebo and its nature as occurring as a result of deception and belief that one is getting a real drug would seem to suggest that non-conscious systems must be centrally involved in the mediation between awareness and the documented physical responses.

Indeed, the Geers et al study cited earlier (Geers, Helfer, et al., 2005) demonstrated that semantic priming (by means of a scrambled sentence task) was an independent predictor of the response to a sleep placebo, which given that semantic priming tends not to be recalled by participants as having an impact on behaviour (Wittenbrink &

Schwarz, 2007b), implies that such priming (and the implicit system more generally) can affect the response to placebos and indeed biologically active treatments. Another study which points in the same direction is the Shiv et al 2005 (Shiv et al., 2005) study which demonstrated an effect of price of an energy drink effecting the number of puzzles solved by participants in a particular time. This effect disappeared when participants attention was drawn to it, which again implies that implicit systems were involved. Despite this evidence, the dominant theoretical framework remains untouched. In this chapter, a new model is proposed that incorporates both of these systems, and makes a number of predictions that will be either supported or rejected by further research.

3.2.3 Embodied Cognition and Placebo

Another issue with most of the conceptions of placebo current in today's research is that they are almost exclusively cognitive, a point made most forcefully by anthropology researchers (Thompson, Ritenbaugh, & Nichter, 2009). This seems strange, given that it rests on assumptions regarding the relationship of mind and body which placebo-related research would seem to argue against. It seems that the theoretical perspective within the field is that the mind may effect the body, but not vice versa. Unfortunately, this is an untenable proposition, as recent work on haptic cognition (where judgements made by participants are affected by the sensory input they are receiving at the time), published in *Science* in 2010 (Ackerman, Nocera, & Bargh, 2010). Indeed, more recent work by Cwir (Cwir, Carr, Walton, & Spencer, 2011) suggests that awareness of one's own interoceptive processes appears to be a good predictor of people's ability to predict the emotional states of others.

There has been a resurgence of interest in mind-brain-body feedback loops within psychology and the social sciences more generally of late. It seems that the dominant cognitive model of mind (a kind of splendid isolation for the brain) is being slowly worn down by experimental evidence. An approach such as this was also proposed by Meissner et al in 2007 (Meissner et al., 2007) in their meta-analysis which showed large placebo effects in some areas and none in others. They suggested that the larger placebo effects may have occurred in some conditions but not in others as the places in which placebo effects occurred tended to be those systems which have large nervous system connections with the brain, as opposed to communication typically mediated through hormones, which are many orders of magnitude slower. One problem with Meissner's theory is that he posits that placebo effects do not occur in clinical trials when the outcome of interest is a hormone such as cortisol.

However, the field of psychoneuroimmunology has noted that psychological variables (specifically optimism) exert major influences on antibody responses (Carver et al., 2010). There is also some emerging evidence that mindfulness based treatments

appear to affect antibody responses, at least in cancer patients (Ledesma & Kumano, 2009). This begs the question of why the analysis by Meisner *et al* found no such effect. One explanation for these conflicting findings would be that the placebo effect is mostly determined by current state effects, while those effects investigated by psycho-neuro-immunology are the result of certain trait like characteristics of individuals.

One psychological state which may have an impact on placebo responses, and should if this theory is to hold is that of mindfulness. Mindfulness is often defined as a moment to moment awareness of somatic and mental processes. If, as the results of Geers *et al* suggest (Geers *et al.*, 2006), somatic focus can increase the size of placebo effects, then mindfulness (or another proxy marker for somatic focus) should also moderate the size of observed placebo responses. If this hypothesis is correct, then mindfulness should be correlated with placebo response, as well as acting as a mediator between implicit and explicit measures of the same construct.

If attention is conceptualised as a limited resource, then placebo effects should be enhanced by placing participants in conditions of low stimulation while the treatment is applied. However, this hypothesis causes problems when considered with the effects of participant provider interaction, which appear to account for a significant portion of observed placebo effects (DiBlasi *et al.*, 2001; Kaptchuk, Kelley, Conboy, *et al.*, 2008).

3.2.4 Additive Models of Placebo

The common approach throughout clinical trials, and the study of placebo effects more generally, is that the effects of drug and placebo are additive. This assumption leads nicely to the principle that placebo effects and drug effects can be separated precisely. However, not all of the evidence points in this direction, and (apart from statistical convenience) there exists no *a priori* reason why this should be the case. Indeed, the work of Geers *et al* on somatic focus would seem to suggest that paying attention to somatic experience can increase the size of placebo effects. This could be occurring because of a feedback loop whereby a treatment is applied, the patient's awareness of the treatment leads them to attend to sensations related to the treatment, which engages the body's own healing systems, which then increase the size of the effect over and above what would have happened without awareness of treatment on the part of the patient. It is perhaps for this reason that drugs which are administered by an automated process are less effective (Benedetti, Rainero, & Pollo, 2003).

Another issue to consider in terms of mathematical models of placebo response is the phenomenon of the active placebo. This is where an active drug is offered as a treatment for which it has no efficacy, and nonetheless this treatment produces larger

healing effects than a typical sugar pill placebo. This theory would argue that this phenomenon occurs because the side-effects of the drug produce a feedback loop whereby the treatment has a somatic impact, which alerts participants to the treatment, causes them to accept it as more credible, and thus activates the body's own healing systems. This process, over time, could easily create a conditioned response to the original non-effective treatment, and loops such as this could be responsible for the observations regarding the efficacy of conditioned placebo responses. In this sense, I am making the argument that conditioned placebo responses are turned into expectancies over time, similarly to the theory of Stewart-Williams (Stewart-Williams & Podd, 2004).

3.2.5 Social Aspects of Placebo

In addition to the possibility of feedback loops between awareness and sensation contributing to the placebo effect, the role of the provider needs to be emphasised also. In most drug research, the treatment itself (the pill or cream) is only one element of the context in which the healing process takes place. In addition to the internal factors which shape response to treatment (optimism and expectancies more generally), there are also important social and environmental factors. For instance, the classic work of Gracely et al (Gracely, Dubner, Deeter, & Wolskee, 1985) demonstrated that the effects of the awareness of the provider can have a large impact on the outcome. In this study, half of the dentists were informed that half of the patients would receive placebo.

In fact, all of them received the real painkiller. However, compared to the group treated by dentists who had not been told that placebo was a possibility, the other patients reported significantly higher pain. Indeed, a systematic review of health care interventions (DiBlasi et al., 2001) provided evidence that provider characteristics accounted for a large proportion of the observed placebo effects. More recently (Kaptchuk, Kelley, Conboy, et al., 2008), a three armed randomised control trial of acupuncture demonstrated that healing rates were greatly increased when the provider spent more time with the patient discussing symptoms and treatment (45 minutes as opposed to 15). This would seem to suggest that one of the reasons so many people use alternative treatments is not the efficacy of the particular form of treatment, but rather the chance to discuss their symptoms in detail and for longer periods of time.

Another important feature of the context in which healing takes place is the social context, that is the shared beliefs and rituals that make up a culture. For instance, Valium has a powerful resonance in our culture, immortalised in songs by the Rolling Stones and glorified through the media. Therefore, even when participants have never experienced the drug itself, they bring pre-conceived notions of what it can do, and

apply these to their perception of treatment, which alters its efficacy. However, research using the open-hidden paradigm (discussed previously) would seem to suggest that Valium lacks any efficacy for anxiety relief when participants are not aware that they are taking it (Benedetti, Rainero, & Pollo, 2003). An additional example of this effect can be seen in the prescription of antibiotics for viral infections. Even though doctors know that they will have no effect, they are still administered to patients. It would be extremely interesting to give people placebo antibiotics and assess its impact on viral infections, relative to the efficacy of true antibiotics. My thesis would suggest that real antibiotics should be slightly more effective, given their obvious side-effects, but that this difference in standardised means would be less than 0.1.

3.2.6 Towards an Embodied Conception of Placebo

Bearing the previous sections in mind, we can now move forward into proposing a new model for placebo, the embodied placebo model. This model, while not rejecting outright the findings of the expectancy theory, aims to enhance it with more recent research. Indeed, this theory is also compatible with that of conditioned placebo, and also with the motivational concordance approach of Hyland et al (Hyland et al., 2007) and the motivational model of Geers (Geers, Helfer, et al., 2005). The essential features of the embodied model of placebo are as follows. Placebo effects are those healing effects which arise from the perception of treatment or caring on the part of the health care provider or from the context surrounding the treatment. They are mediated by four different kinds of factors.

- They are mediated by implicit cognitions operating extremely quickly
- They are also impacted by awareness of the treatment and cognitive models of its efficacy
- In addition, these implicit and explicit attitudes are enhanced or denigrated by somatic sensations
- They are enhanced or reduced by the communications exchanged between the participant and the provider
- They are also affected by the system of communications surrounding both the participant and the provider
- Finally, they are mediated by somatic feedback from the surrounding environment.

In an effort to advance research, this theory makes a number of key predictions:

- First, that placebo effects will be correlated with scores on implicit measures

- Second, that there will be an interaction effect between explicit and implicit attitudes on placebo response
- Third, that increased interoceptive awareness (or mindfulness) will increase the size of placebo responses
- Fourth, that the implicit attitudes of the practitioner will exert a significant impact on the placebo response of the patient ¹
- Fifth, that adding sensory input to a placebo treatment will increase its effectiveness.
- Sixth, that the extent of changes in physiological measures in participants will be correlated (as a lagged variable) with the size of the next reported pain rating.

However, in order to properly test this model, it needs to be compared to other plausible models. The first of these is the model of Kirsch, noted in his 1985 paper. This model claims that expectancies are the major (indeed only) mediating factor between consciousness and placebo responses. Therefore, this model suggests that there is a direct effect of expectancies on placebo response, that physiological outcomes should not be predictive of the placebo response and that any effect of optimism and mindfulness will be mediated through expectancies. This model claims that there will be no direct effects of any other explicit or implicit measure on placebo response, but that they will all be mediated by expectancies.

Another, equally plausible model is that optimism is the driver of the placebo response, and that the effects of expectancies are mediated by levels of optimism. This model would suggest that both implicit and explicit optimism will have direct effects on the placebo response, and that expectancies (implicit and explicit) shall be mediated by levels of optimism. This model makes no predictions about the effects between physiological response variables and placebo response.

Additionally, a variation of each of these models would suggest that either explicit or implicit expectancies could be the sole mediator of the placebo response, and these models were also tested.

Finally, a null model where none of the variables measured had a relationship to the placebo response was examined, to ensure that the other models were at least somewhat more plausible.

These models will be examined using a structural equation modelling approach, which should allow for an efficient and accurate test of this theory, even if the population is not particularly representative.

¹due to researcher constraints, this hypothesis was not tested in this thesis

3.3 Development of the IAT

Qualitative analysis was an essential part of this project, and the details are reported in Appendix B. It gave insight and data into the development of the IAT's used in the final part of the research project. The methods used for qualitative analysis here involved a thematic analysis (Braun & Clarke, 2006) of the interviews conducted was carried out to develop themes both for the repertory grid and for the IAT.

Qualitative research typically relates to the analysis of interviews and other texts derived from people. It differs fundamentally from quantitative analysis in that it aims for a deep understanding of particular individuals, while quantitative analysis aims for a broad understanding of the sample as a whole.

The biggest problem with qualitative analysis is that it cannot scale to the level of large scale surveys, as it requires significant amounts of researcher time per participant while quantitative surveys have a cost of development in time, but the marginal cost of administering the survey to a new participant is essentially zero (assuming distribution over the internet).

The issue of reflexivity is crucial to qualitative research (and also appears in quantitative research, though rarely as openly) (Rosenthal, 1967, 1969).

Reflexivity refers to the impact of the researcher's prior conceptions and approaches have on the course of the interviews (Finlay, 2002).

This is extremely obvious in the choice of the major questions to be asked in the interviews, but it can occur in subtle ways during the interviews also (for example in the use of language by the interviewer) and in the quality of communication or rapport experienced by the researcher in the course of the interview. Reflexivity is also critical during the analysis, as the researcher must be aware of their own biases and ensure that this affects the analysis as little as possible, or at least report where the problems arose for them. A process of thematic analysis was used in preliminary interview research for this thesis, and the output of this was used to inform the development of the repertory grid.

3.3.1 Repertory Grids

The use of repertory grids in this research was as a bridge between the qualitative analysis carried out and the quantitative side of the research.

Repertory grids were developed by George Kelly as an aid to therapy (G. Kelly, 2003) although it has been used in many diverse situations in the ensuing years. Repertory grids were developed out of Kelly's theory of cognitive consistency, an active area of

research which fell out of favour following the demonstration of cognitive dissonance by Festinger in 1947 (Greenwald et al., 2002).

The premise of the technique is simple. Firstly, participants are supplied with a list of important people in their life, such as their mother, an older sibling and a teacher whom they liked or disliked. They write down the names they have chosen for each person, and then they compare the people in groups of 3. For each group (or sort), they are asked to describe how two of them are similar and also how one of them is different in a word or short phrase. These words or phrases can then be analysed both quantitatively or qualitatively.

The primary method of quantitative analysis was through factor analysis.

The approach taken in this project was as follows. In one of the surveys carried out on the UCC population (TCQ version 1) participants were asked to rank the most important people in their life who were related to health care. This data was then sorted and ranked, and a list of the most common people used was compiled into a health related repertory grid. This was then administered to a small sample (N=17) to test the instrument. The results of this testing are described in Chapter 6, in Section 6.4.

3.3.2 Development of IAT

The plan for the development of the IAT was to use the constructs obtained from the repertory grids to develop useful stimuli for the IAT. Unfortunately, as Chapter 6 describes, this portion of the research did not lead to a successful outcome, for reasons described in that chapter.

Therefore the IAT was developed from both the important figures which arose from the repertory grid, the qualitative interviews (see Appendix B), and to match the explicit measure of treatment credibility. The optimism IAT was developed along similar lines to most other Implicit Association tests, in that the survey for this measure was used as a base.

3.4 Quantitative Research Methodology

This section will describe the methods employed for each part of the thesis and provide a rationale for why these methods were used in the thesis. A mostly quantitative approach was taken for the following reasons. Firstly, the placebo is a very noisy phenomenon (Singer, Grossman, Avidan, Beckmann, & Pe'er, 2005), subject to many sources of error and bias (c.f. Section 2.2). Secondly, in order to

predict the placebo effect, there needs to be some kind of measure, and these are most appropriately conceptualised in numerical terms

This thesis consisted of three main parts. Firstly, following a thorough literature review, the major constructs associated with the placebo effect and implicit measures were identified. These constructs were then administered to large samples of the population from which the experimental participants were drawn. This procedure was carried out for two reasons. In the first case, this was so that the population means could be estimated more precisely and thus the experimental sample compared on these measures.

The second reason was so that more sophisticated models could be developed for person responses (which typically require larger samples than are common to experimental studies) and could then be applied to the experimental sample. This approach marries two of the strengths of psychological research; the latent variable approach common in psychometric research; and the use of rigorous experimental design to determine theoretical relationships between constructs (through Structural Equation Modelling). This thesis aimed to use both of these strengths in combination to gain insight into the causes underlying the response to placebo in healthy volunteers.

The second major part of the thesis was the development of the implicit association tests (IAT's) and the explicit measure of expectancies (treatment credibility questionnaire) used in the experimental portion of the research.

The third, and final, part of the thesis was the testing of these measures in an experimental setting using a placebo analgesia design examining the response to ischemic pain in healthy volunteers.

In this section, the statistical techniques used to answer the primary research questions shall be discussed. Only methodology used throughout the thesis will be described in this section. Therefore, methods of cross-validation will be described, as will the regression models utilised and finally Structural Equation Modeling shall be discussed. The methods employed only within one chapter are discussed in the appropriate chapter.

3.4.1 Problems of Sample Inference

3.4.1.1 The Problem

In every statistical approach, the core is the development of inferential tools to reduce our uncertainty about the events under study (Gelman & Shalizi, 2010). Given that we typically lack infinite resources, sampling from populations in a randomised manner is used to approximate the quantities of interest (Venables & Ripley, 2002).

However, the specific sample we have recruited is rarely interesting; it is wanted to infer properties of the population from which they are drawn. In non-technical terms, given a sample and some analyses, a model is developed, which is hoped will predict the behaviour of future samples (and indeed the population).

This approach toward inference is often operationalised in the creation of a model, whether based on the results of a linear regression or factor analysis. Typically, the aim is to maximise the amount of the response variable in a sample of size n explained given some number of parameters. It is trivial to see that as the number of parameters (p) increases, so does the fit — in the limit, this would involve the fitting of a model with a separate parameter for each observation.

Clearly, such a model will violate the principles of parsimony and clarity that we aim for in our science. However, even when p is less than n , we still run the risk of over fitting a model to our data. Over fitting is said to have occurred when features of the data that are essentially random are modeled (Friedman, Tibshirani, & Hastie, 2009).

Because factor analysis is lenient towards mis-specified models and tends to model error as well as signal, many psychometric theories have faltered on the rock of replication (Fabrigar, Wegener, MacCallum, & Strahan, 1999). SEM is often used as a panacea for such problems. However, a Structural Equation Model is only as good as the data and theory behind it, and if the factor analysis models noise, so too will an SEM on the same data set (to a lesser extent, of course).

3.4.1.2 A Solution

The typical scientific approach to this problem is simple — replication. Replication, preferably by independent researchers, is supposed to ensure that models eventually tend towards the minimum of parameters for the maximum of explanatory power.

Indeed, many fit indices penalise complex models over simple models. This suffices for some research. Replication can also be fraught with difficulties, as it takes time, effort and additionally makes the assumption that population quantities are stable over time. Nonetheless, it is the ideal solution. However, given that replication is not incentivised by the scientific community (and may be unethical in some situations) (Roediger, 2013), researchers in the field of machine learning have come up with a novel approach which appears to improve predictive accuracy and can also aid in the development of theoretical understanding (Friedman et al., 2009).

3.4.1.3 Cross-Validation

The solution proposed by those machine learning researchers is at once simple and elegant, while also extremely practical. The technique is known as cross-validation,

and its application routine in commercial and scientific data-mining settings. However, it does not appear to have found much favour within psychology as of yet.²

The basic premises of the techniques are:

- All models are wrong;
- Models are best tested on independent samples;
- Independent samples are sometimes hard to come by;
- Therefore, data sets should be split into training and test sets, where the model is developed on the training set, and its accuracy assessed on the test set.

This approach seems to improve predictive accuracy by an order of magnitude, especially when applied to large data sets (Breiman, 2001).

The principle of training and test sets has since been generalised to k -fold cross validation where the data set is split into k random pieces, and all but one of these are used to estimate a model, while the other is used as a test. This procedure is repeated k times, and the results are averaged to form the best model (for that sample of data, at least). Some authorities argue that this procedure should be repeated at least k more times, to control for the effects of random sampling (Friedman et al., 2009).

Another variation on the central approach is leave-one-out cross validation, where given a sample of n observations, fit a model on $n - 1$ and test on the other, a total of n times. This approach, while taking the technique to its logical extreme is not of particular usefulness to this research, as large inter-personal variability between individual participants typically observed in psychological data would tend to reduce its efficacy (Friedman et al., 2009).

In essence, cross-validation is an extremely valuable technique which has been mostly ignored in psychology. It is the opinion of this researcher that this technique is useful, and it will be applied consistently to this research.

3.4.1.4 Regression Models

At the heart of typical psychological modelling practice lies the general linear model. This model underlies such familiar techniques as correlation, regression and analysis of variance (ANOVA) (Gelman & Hill, 2007). These techniques are based on the idea of fitting a straight line (or plane, in the case of multiple predictor variables) to the observed data and using this line to make inferences about the relationships between variables of interest.

²with some notable exceptions (Dawes, 1979)

The general linear model has been elaborated greatly over the last century, and has been applied to the approximation of relationships that do not meet all of the assumptions noted above (Gelman & Hill, 2007). The introduction of link functions (non linear functions designed to transform the response variable into a form suitable for the model) led to the development of generalised linear models, which allow the same computational techniques to be used to fit and test models for data which does not fit the requirements of the standard linear model (McCullagh & Nelder, 1989).

For example, logistic regression is a technique for prediction of binary variables using a link function. Poisson regression is used to model data which is bounded by zero and positive infinity. A number of quasi methods are available for both of these techniques which allow models to be fitted to data which has an excess of zeros (so called zero inflated models) (Gelman & Hill, 2007; Venables & Ripley, 2002). In another direction, the requirement for the residuals to be uncorrelated has been relaxed to allow for the development of multilevel or mixed models, which allow the residuals of particular groups to be correlated with one another (Gelman & Hill, 2007).

A number of major issues arise with the use of the general linear model with psychological data. Firstly, the requirement of normal errors can sometimes be difficult to satisfy. This follows from the manner in which the normal distribution appears to behave. When it was originally discovered (by Gauss) (Stigler, 1986) it was used to model the combination of many small, independent variables. It tends to work well as an approximation when there are many independent variables affecting the results of an analysis.

However, when psychological tests (such as self report instruments) are developed, the aim is to remove as many of these small influences as possible so that the measure taps one construct with clarity. This will often lead to a non-normal distribution of errors. Non parametric approaches are an alternative to the General Linear Model, but these often lack the power of their parametric alternatives. The central limit theorem assures us that, in the limit, any distribution of means will converge to a normal, and in practice perhaps as little as 100 observations may suffice for a t-test (Venables & Ripley, 2002).

A more serious problem (in terms of the impact on outcomes) is heteroscedasity, where the response variables (or the predictor) do not possess constant variance across their range (Gelman & Hill, 2007). This can seriously affect the models built as points with less variance will be much more influential than those with a higher variance. The assumption of homoscedasity can be checked by formal statistical tests, but often graphical tests of this assumption are much more revealing, as they can show where the model fails as well as whether or not it fails.

In this research, linear and logistic regression models were utilised throughout all of the quantitative research in order to test hypotheses.

3.4.2 Psychometric Analyses

The kinds of psychometric models employed in this thesis were threefold, Factor Analysis (FA), Item Response Theory (IRT) and Structural Equation Modeling (SEM). The general approach taken was to use FA and IRT to develop models, and utilise SEM techniques to test these proposed structures. Below, these three methodologies are introduced, following which the general plan of analysis for Chapters 4 and 5 are described.

3.4.2.1 Factor Analysis

Factor analysis has a long history in psychology, and is now over one hundred years old. It is the most commonly used latent variable modelling technique in psychology, and more pages of *Psychometrika* have been devoted to it than to any other technique (Henson & Roberts, 2006). Despite this, there are still a number of issues and controversies which surround the technique (Sass & Schmitt, 2010). Essentially, factor analysis is an attempt to approximate a correlation matrix with a smaller matrix of parameters. These hypothesised latent variables tend to be called factors or components. As such, it is a dimension reduction technique, much like principal components or clustering.

The most critical issue surrounding factor analysis concerns determination of the number of factors to extract (Zwick & Velicer, 1986). This is an important issue, as while other criteria such as rotation can have an impact on the interpretation of the issue, the decision to extract too many or too few factors can prevent any real understanding of the foundations of a scale.

The issue is not that there are no criteria on which to base a principled decision, but rather that the different criteria often do not agree, and it is thus ultimately left to the informed opinion of the researcher which factor solution is to be preferred. All of the decision criteria will be reviewed in turn, and their advantages and disadvantages will be discussed (Henson & Roberts, 2006).

"Solving the number of factors problem is easy, I do it everyday before breakfast.
But knowing the right solution is harder" (Kaiser, 1954).

The choice of criterion for retention of factors is extremely important in applied work. This is because if an incorrect number of factors are extracted, then the predictions for the experimental portion of the research will be biased, and thus will not prove as useful as the method could otherwise be.

The first, and most popular, criterion is surprisingly the least useful (Zwick & Velicer, 1986). This rule is called eigenvalues greater than one criterion and recommends keeping all factors whose eigenvalues are greater than one. The rationale behind this approach is that eigenvalues less than one explain less of the variance in the matrix than one item, and as such should not be retained. More recent research appears to put the minimal criterion for retention of eigenvalues at approximately 0.7 (Henson & Roberts, 2006).

The second criterion often used is the scree plot technique, which was popularised by Raymond Cattell. This criterion recommends that the eigenvalues of all factors should be plotted against their number, and only factors before the drop off in eigenvalues should be used. As the process of factor analysis ensures that the first factor will have the largest eigenvalue, followed by the second and so forth, this criterion looks for the point where the eigenvalues are very close to one another. This criterion has a number of advantages. It is available in all statistical packages, it can be used without any special training and it tends to give results which are somewhat, if not totally accurate (Zwick & Velicer, 1986). Its major disadvantage is that it relies upon the interpretation of the researcher, but given the strong emphasis on interpretation throughout factor analytic literature this should not be regarded as too much of a handicap.

The next criterion which can be used is that of parallel analysis. Parallel analysis is a Monte Carlo (simulation) technique which simulates a data matrix of equal size and shape to the matrix under study, and calculates the eigenvalues of these simulated matrix against those of the real matrix (Horn, 1965). All factors are retained up to the point where the simulated eigenvalues are greater than the true eigenvalues. Parallel analysis is one of the better techniques for assessment of the number of factors to extract, and it can often give very accurate results (Zwick & Velicer, 1986).

Its major disadvantage is that it tends not to be available in many statistical packages, and that it can often over factor the data-set. Additionally, many tools simulate the new data from a normal distribution (Micceri, 1989). It does produce some of the most accurate results in simulation studies so it is a useful tool in practice (Zwick & Velicer, 1986).

Another useful criterion is that of the Minimum Average Partial Criterion (MAP) which extracts factors from the data-set until only random variance is extracted (Revelle & Rocklin, 1979). Again, this is an accurate criterion (Zwick & Velicer, 1986) which is little used as it is not available in popular statistical programs. The only problem that has been found with this criterion is that it tends to under-extract factors.

However, the ultimate test of a factor solution (without using other methodologies, such as Structural Equation Modelling) (Joreskog, 1978) is its theoretical clarity and

interpretability, and this will be the first test used for all proposed factor solutions.

Another area of dispute amongst researchers in the factor analytic field is which method of rotation to use (Sass & Schmitt, 2010). As the eigenvalues are only defined up to an arbitrary constant, these rotations do not have any substantive impact on the factor matrix, except that they can make it easier to interpret (which is normally very useful).

Rotations are commonly applied to factor solutions in order to reduce items loading on multiple factors, and to aid in the discovery of simple structure (Henson & Roberts, 2006). Rotations can be divided into two classes, orthogonal and oblique (Sass & Schmitt, 2010). Orthogonal rotations return uncorrelated factors, while oblique rotations allow the factors to be correlated. Given that most psychological measures are correlated with one another, one would expect oblique rotations to be more common. However, the default appears to be orthogonal rotations, as they are claimed to be easier to interpret (Henson & Roberts, 2006). Oblique rotations were applied throughout this research, as if the factors are truly uncorrelated, then the oblique rotation will show that, while the converse is not true for orthogonal rotations.

In conclusion, factor analysis is an extremely useful technique which has been widely applied in psychology. It is available in most software packages, is typically easy to interpret and can be carried out with a small number of items (more than 200 is often sufficient). The major problems with factor analysis are the necessity of determining how many factors to extract, which is a difficult and often subjective decision, and additionally if scores are required then many methods exist, again without any clear rationale for choosing one over the other.

3.4.2.2 Item Response Theory

Item response theory (IRT) is often called model-based measurement (Fischer & Molenaar, 1995) and is a newer approach to analysing self report data, developed both by the Danish mathematician Rasch in work for the Danish army, and also separately by Lord and Novick in the US, while working for the Educational Testing Service (ETS) in the 1950's and 60's (Van der Linden & Hambleton, 1997). The fundamental premise of IRT is that the properties and scores on a psychometric test can be modelled as functions of both the items on the test and the "ability" of the people taking the test.

Note that despite their similarity, the assumptions and methods used in Rasch modelling and IRT modelling are actually quite different. Rasch models require the sum score of the answers to the questions to be a sufficient statistic for the latent trait. This requirement allows for the decomposition of item scores from person scores, and is core to the assumptions (and provides a useful set of fit-statistics, such

as Infit and Outfit Mean Square measures). IRT (as developed by Lord & Novick) does not assume that the sum score is a sufficient statistic for the latent trait. This means that different items can have different contributions to the latent trait, and means that easy separability of person and item parameters is much more difficult, and additionally that fit statistics are more difficult to interpret and justify.

The IRT approach suggests that conditional on both the ability of the person and the difficulty of the test, the responses of each participant can be predicted probabilistically. As the latent ability of the participant rises, they tend to choose alternative responses which are more reflective of this latent ability. One example of this might be an item for extraversion “I am always the life and soul of the party“, those respondents who had a higher latent score on the extraversion construct would tend to choose the agree or strongly agree options (on a typical five point scale).

For instance, if one was modelling extraversion using a set of ten items, the participants who scored highest in extraversion would be most likely to respond strongly agree to the items. IRT also assumes a property called local independence, which states that conditional on the ability measured by the test, the scores of each participant are independent of one another. This assumption was checked throughout the analysis in the following chapters. Additionally, another assumption of these models is (latent) monotonicity, in that if a participant responds Neutral to a particular item, then a participant who responds Slightly Agree should have a greater latent trait than does the first. Again, this assumption was checked throughout the analysis.

There are a number of different approaches taken to IRT (Van der Linden & Hambleton, 1997; Fischer & Molenaar, 1995). The Rasch models are the simplest, and have a number of extremely appealing mathematical properties. These models assume that only one trait is measured by the items, that all items are equally predictive of the trait, and that there is no guessing (Van der Linden & Hambleton, 1997).

Because of these assumptions, it is possible to separate out person abilities and item difficulties perfectly. However, another approach (normally referred to as a two parameter model, or IRT proper) claims that items are differentially predictive of the ability being measured, in a manner analogous to different strength of loadings of items on a construct in factor analysis. Another model the three parameter model (Lord, Novick, & Birnbaum, 1968), allows for correct responses through a process of guessing, but this model is not normally applied to polytomous items (Van der Linden & Hambleton, 1997; Mair, Hatzinger, & Maier, 2010).

In general, IRT models are represented by the logistic function, and are estimated iteratively through procedures of numerical optimisation (maximum likelihood (Fischer & Molenaar, 1995)). The function used to describe the data is a logistic one, where ability is estimated from the probability of answering the question correctly.

The difficulty of an item is conventionally defined as the ability of participants who answer the question with 50% accuracy. The parameter *beta* is defined as the difficulty of the item. In two parameter models, another parameter α is defined and is used for the discrimination of the item (the slope of the curve). In the more complex three parameter model, θ is used to measure guessing (the probability of a correct answer given low ability) (Van der Linden & Hambleton, 1997).

IRT was developed in the context of ability tests, and this leads to much of the vocabulary fitting uneasily within personality psychology. For instance, in the context of a credibility questionnaire about various treatments, the questions on homeopathy are categorised as most difficult (see Chapter 5). This does not mean that they are harder to answer, just that the probability of a respondent endorsing them is lower than the probability of a participant endorsing a similar item on the efficacy of painkilling pills.

It is important to note that the names of the models are slightly deceiving, while they are called 1, 2 & 3 parameter models, they actually involve the estimation of 1, 2, or 3 parameters for each item. This normally means that a test of $1, 2, 3 \dots, n$ items will require the estimation of either $n, 2n, 3n$ parameters.

3.4.2.3 Structural equation modelling

Structural equation modelling is regarded by many as an adjunct technique for evaluating the results of particular factor solutions (Fabrigar et al., 1999). however, it is actually a set of far more general techniques to test the relationships between both manifest and latent variables, and even to establish causality in some cases (Pearl, 1998). the factor analytic procedure is full of interpretative procedures where no principled choice can be made, and structural equation modelling (hereafter SEM) is an attempt to compensate for some of these deficiencies.

SEM was developed by Joreskog in the 1970's (Joreskog, 1978). it provides a means of testing hypothesised relationships between latent and manifest variables. in practice, the result of a factor analysis is regarded as a measurement model of the data.

This is combined with a structural model (which describes how the latent variables relate to one another and to the manifest variables). The two of these models are then used to construct a covariance matrix which is then compared with the observed data, and a number of indices of model misfit are calculated. foremost among these is the χ^2 , which estimates the degree of model misfit. the desired result is a p-value of greater than 0.05, which shows that the two matrices are not significantly different.

However, the χ^2 is extremely sensitive to sample size, and tends to be rejected in almost every case (Henson & Roberts, 2006), given that the sample sizes needed for accurate factor analysis and structural equation modelling tend to be quite large. as a

result of this, many other fit indices have been developed. foremost amongst these are the non normed fit index (NNFI), which is also known as the tucker lewis index (Bentler, 1990), the root mean square error of approximation (RMSEA) (Rigdon, 1996) and the bayesian information criterion (BIC) (G. Schwarz, 1978) and an information criterion (AIC) (Akaike, 1974). These all have different strengths and weaknesses and are typically used in a complementary way. All of these fit indices incorporate explicit penalisation, which aids in avoiding overly-complex models. the χ^2 it also has some penalisation (based on degrees of freedom) but it is typically not strong enough to prevent over-fitting. some authors argue that this focus on other fit measures apart from χ^2 is a way to avoid rejecting favoured models, but such a view is controversial in the field at present (Barrett, 2007).

The multivariate normality assumption made by the procedure is often difficult to meet in practice. however, there are a number of distribution free methods in SEM, of which the most common is a weighted least squares approach. this proceeds similarly to a weighted least squares approach in linear regression, where points are assigned weights depending on how closely they meet the assumptions of the model. sample size, by contrast, is typically easy to increase (at least for non-clinical populations).

Identification of the model is one issue in practice, though as Joreskog notes, this can often be achieved by fixing a number of parameters to 0 or 1 (the inter-factor variances are often scaled in this fashion) (Joreskog, 1978).

Another, more theoretical issue is that no set of data is uniquely determined by an sem model. this is known as the problem of rotation in factor analysis (MacCallum & Austin, 2000). given a covariance matrix w , and a set of data d , there are many solutions which provide the same fit indices of the model to the data. this can lead to a similar problem as occurs to factor analysis, where the researcher must make a choice between models which are quantitatively identical. one approach for resolving this problem was discussed above, in Section 3.4.1.

In this thesis, SEM was applied extensively to test the theorised relationships between variables, and the experimental chapter includes a number of tests of theoretically derived models (c.f. Chapter 7).

3.4.3 General Analytical Approach

Factor solutions were extracted using principal axis methods from the psych R package (Revelle, 2015).

primarily, direct oblimin methods of rotations were utilised, but promax rotations were also applied where these failed to converge. In general, promax will attempt to find the rotations which maximise simple structure.

After the various factor structures were obtained, they were plotted and analysed for interpretability. Communalities and uniquenesses were assessed to ensure that there was no over or under factoring in the solutions³.

Following this procedure of extraction and interpretation, Structural Equation Modelling was applied to each of the proposed factor solutions using the OpenMx (Boker et al., in press) and lavaan (Rosseel, 2012) packages for R. The optimal factor solution was chosen using the AIC of each fitted model, along with the RMSEA of the proposed solutions, and these solutions were then evaluated in terms of their performance on unseen data.

Following the investigation of structure with the methods of classical test theory, the scales were analysed using Rasch models and item response theory (using functions from the ltm package (Rizopoulos, 2006)). Firstly, Mokken analyses (from the Mokken package (Van Der Ark, 2007)) were run, in order to check the assumptions of monotonicity, local independence and to assess how many sub-scales the analysis should be carried out on.

Following this, two successively more complicated IRT models were fitted to each sub-scale (one parameter with a single α for all items and two parameter, with a separately estimated α for each item). Graded Response Models were used for IRT estimation as these are appropriate for ordinal data. Additionally, the most successful IRT models were used as the basis for other models analysed using SEM.

The IRT models were cross-validated using the following procedure. Firstly, the factor scores for an appropriate GRM (1 parameter or 2 parameter) were fit to the data. Then, the most successful model from the previous data was used to produce factor score predictions. The square root of this difference was then taken, and a penalisation parameter of $\ln K$ was added, where K was the number of parameters in the model. Additionally, the Pearson correlation coefficient was calculated between the two factor score outputs. The IRT models were then ranked using this number, where lower numbers represented a better fit to the unseen data.

Linear regressions were run to examine the differential effects of each of the correlated variables. Step wise selection on the training set was carried out, but the p-values are always reported from the test set. The performance of each of these methods was then assessed on held-out data (from another split or sample, using ten fold cross-validation as previously described).

In the case of the first samples in Chapters 4 and 5, some of the second sample was used as a held out data set. For the second sample, the entire data set was split into three or four splits, and the cross validation procedure carried out for each. The splits

³Communalities are important as they represent the common variance actually included in the model, in contrast to PCA which uses all variance in the matrix

were kept quite large to allow for psychometric models to be fit to each split separately, and then to be tested on the remaining data. In Chapter 4 a back testing approach was used, whereby the most successful models from the second sample were assessed for their performance on the first sample. For Chapter 5, this was not possible due to the revision of the instrument, and so the final model was selected on the basis of the entire data-set.

3.5 Conclusions

In this chapter, the theory underlying the work of this thesis has been explained, and the methods used to answer the primary research questions have been addressed. Additionally, due to the complexity of many psychometric models, cross-validation techniques were employed to provide unbiased estimates of their likelihood. Structural Equation Modeling was employed pervasively throughout the thesis to test particular factor analytic and regression model structures.

Chapter 4

Health Optimism and Mindfulness Data

4.1 Introduction

This chapter reports an investigation of the psychometric properties of three measures, the RAND Medical Outcomes Survey (RAND-MOS), a measure of self reported health, the Mindful Attention Awareness Scale (MAAS), a measure of mindfulness and the Life Orientation Test, Revised (LOT-R), a measure of optimism.

The major experimental work of this thesis involved the placebo effect, and as the research considered the utility of a new measure to predict this response, it was important to administer some instruments which have been shown to be associated with it in previous studies.

Additionally, as the new method proposed to predict the placebo response in healthy volunteers involved the use of implicit measures (two IAT's) (c.f. Chapter 2) it was also necessary to collect data on the construct of mindfulness as operationalised using the Mindful Attention Awareness Scale (MAAS), as it has been shown to mediate the relationship between explicit and implicit measures (Levesque & Brown, 2007).

Over the course of the study of the placebo, there have been few individual-level psychological predictors of the effect which have been reproducible. Expectancy is the most commonly measured construct within the field, and some meta-analytic evidence suggests that it is only significantly associated with the response around 60% of the time (DiBlasi et al., 2001).

Some authors claim that the search is fruitless (Shapiro & Shapiro, 1997) while others argue that the placebo effect is shaped more by the situation, and therefore there is no such thing as a placebo responder (Kaptchuk, Kelley, Deykin, et al., 2008).

However, in recent years optimism (Geers, Helfer, et al., 2005; Morton et al., 2009) has been shown to be associated with the placebo response. More specifically, the Life Orientation Test, Revised (LOT-R) has been shown in a number of studies to be associated with the response to placebo.

Mindfulness is a construct that has both been associated with health (Carmody, Reed, Kristeller, & Merriam, 2008) and moderation of the relationship between explicit and implicit measures (Levesque & Brown, 2007). As such, it is a construct which has been associated with the IAT in previous work (c.f. Chapter 2) and could plausibly be associated with health. Some more in-depth discussions of this potential relationship were outlined in Chapter 3, in Section 3.2.3.

Health (in the form of the RAND-MOS) was included in this piece of research both to replicate the optimism-health link and examine the relationship between optimism and mindfulness.

Additionally, the relationship between these three variables (health, optimism and

mindfulness) can provide some insight into health cognitions, which are theoretically linked to placebo-related cognitions. This follows as if the placebo effect is the result of expectancies, then there should be some shared variance between the response to placebo and other health-related cognitions.

The primary aim of this study was to investigate the relationships between optimism and mindfulness in the population under study, and to provide a basis for the building of psychometric models which could be used to predict the relationship of these variables to the placebo response in the experimental study.

The study formed an opportunity to collect background data for the population of interest, to assess if the participants in the experiment were systematically different from those who had responded to a survey invitation. This is critical if the results from the experimental portion of the research are to generalise to any further samples, given that inferences cannot be made about the experimental sample if it is not understood how they stand in relation to other samples from the overall population which was used throughout the research. This process should allow for more accurate predictions of participant responses, assuming that the models generalise to the new sample.

4.2 Optimism and Placebo

Dispositional optimism is often defined as *generalised outcome expectancies about the future* (Carver et al., 2010). When using this definition, it seems relatively likely that there may be a relationship between optimism and placebo response, and yet this has only been investigated in recent years.

Dispositional optimism appears to exert some influence on placebo effects, in some situations (Geers, Helfer, et al., 2005; Morton et al., 2009). The effect seems to be that those higher in optimism respond better to positive suggestions, while those higher in pessimism respond better to negative suggestions.

Another study found that general (but not specific) expectancies had a significant impact on the response to placebo in a meta-analysis of randomised controlled trials of chronic back pain (Myers et al., 2008). Generalised expectancies are how optimism is defined, given that all expectancies around future states of health are essentially outcome expectancies.

Other studies appear to show that this optimism effect is not general, but rather depends on the context in which the experiment takes place (Hyland et al., 2006). In the Hyland *et al* experimental study, spirituality rather than optimism was a predictor of the response.

Crucially, this only occurred when the treatment was classified as spiritual. When a gratitude based treatment was used, gratitude acted as a predictor. These results suggest that any trait which predicts placebo response will likely only be effective in certain situational settings (Kaptchuk, Kelley, Deykin, et al., 2008).

This would seem to suggest that although there appear to be some dispositional predictors of the placebo response, they are moderated by the context in which the study or treatment takes place.

Alternatively, one could argue that expectancies drive these effects by mediating the impact of other contextually relevant variables. However, to take this position would require that the theory of Kirsch, that expectancies exert direct physiological effects, would need to be abandoned (Kirsch, 1985), and indeed this proposition was tested in the experimental portion of the research (c.f. Chapter 7)

A recent study (Morton et al., 2009) in a placebo analgesia paradigm argues for a stronger interpretation of the role of optimism. This experimental study used a repeated measures design, and utilised a preconditioning method in the first session which is known to increase the size of the placebo response (Voudouris et al., 1985).

While in the first session there was no effect of optimism on the results, in the second study dispositional optimism was significantly correlated with placebo analgesia, explaining 55% of the variance. This would suggest that while optimism may not produce a placebo response in itself, once a response has been produced it can be effective in maintaining it over time.

Hyland suggests that the optimism effects on placebo response are mediated through expectancies, and that when these are not a factor, neither is optimism. This sounds plausible, but the relationship could easily go the other way in that optimism could drive the observed effects of expectancies. This is not a question which can be answered without further empirical investigations, some of which were conducted as part of this research (see Chapter 7, especially Section 7.3.11).

The relationship between optimism and health is well known. Optimism was recently reviewed in relation to its effect on physical health and well-being (Carver et al., 2010), optimism has been used as a predictor for many years in the area of Psychoneuroimmunology (PNI) and appears to be associated with better health outcomes than is pessimism (Baker, 2007; Conway, Magai, Springer, & Jones, 2008). Optimism has also been associated with lower mortality risk in a large longitudinal cross-sectional study of individuals at risk for cardiovascular disease (the Women's Health Initiative) (Tindle et al., 2009).

A meta-analysis has also confirmed this link between optimism and better coping styles, as well as a strong negative relationship between optimism and negative affect (Andersson, 1996; Nes & Segerstrom, 2006). Higher levels of optimism have also

been associated with quicker recovery from surgery, insulin therapy and chemotherapy (Allison, Guichard, & Gilain, 2000). A prospective study looking at outcomes from a group of head and neck cancer patients found that optimists consistently reported better health outcomes than pessimists (Allison et al., 2000).

The question of the mechanism by which optimism manifests differences in health is still unclear. Some researchers argue for a direct effect of optimism on immune function, while others argue that optimism exerts its protective effects through the effects of persistent striving after health goals (Segerstrom, Castaneda, & Spencer, 2003).

4.2.1 Factorial Structure of the LOT-R

The factor structure of the LOT-R (and indeed, the dimensionality of the optimism construct more generally), has been a subject of some debate. The originators argue for a one-factor structure (Carver et al., 2010; Scheier et al., 1994), while other researchers have found that the scale is better modeled as having both optimism and pessimism components. The Geers work referenced above took a latter approach, dichotomising the scores on the scale to create groups of optimists and pessimists¹. Therefore, both one and two factor models were applied to the scale to determine if a one or two factor structure fit better in the population of interest.

4.3 Mindfulness

The construct of Mindfulness or 'attentional control' has been defined as:

'a mental ability which facilitates a direct and immediate perception of the present moment with non-judgemental awareness'

(Kohls, Sauer, & Walach, 2009) (p. 2). Derived from the Buddhist contemplative traditions, mindfulness represents attention to the thought process, rather than to thought content (Brown, Ryan, & Creswell, 2007). Jon Kabat-Zinn popularised the practice of mindfulness in the West (Kabat-Zinn, 2003), developing an eight-week long Mindfulness Based Stress Reduction (MBSR) programme which is now often used in clinical settings for patients with chronic illnesses.

The usefulness of mindfulness in this research was twofold

1. It has been shown to be a predictor of health, and treatment programs developed around mindfulness practice have been associated with improved physiological and psychological outcomes

¹this approach is not without its problems

2. Additionally, the construct has also been implicated as a moderator of the relationship between explicit and implicit measures (Levesque & Brown, 2007)

The combination of these two features was why mindfulness (specifically the Mindful Attention Awareness Scale) was chosen as the measure to test the mediation relationship between explicit and implicit expectancies.

4.3.1 Health benefits of Mindfulness

In a study of cancer patients mindfulness training was found to reduce stress and to improve quality of life (Carlson, Speca, Faris, & Patel, 2007), while in a study of patients who were HIV positive, MBSR training was associated with improved natural killer cell activity (Robinson, Mathews, & Witek-Janusek, 2003). Meta-analyses have suggested that MBSR programmes have an effect size of $d = 0.5$ for psychological variables such as quality of life and mental health, and $d = 0.2$ for physical outcome variables such as cortisol levels and immune function (Grossman, Niemann, Schmidt, & Walach, 2004). Measures of the construct of mindfulness date from the early years of this century. Mindfulness has been operationalised into a number of different scales including the Mindful Attention Awareness Scale (MAAS) (Brown & Ryan, 2003), the Kentucky Inventory of Mindfulness Skills (Ruth, Gregory, & Kristin, 2004) and the Five Facet Mindfulness Questionnaire (FFMQ), which was developed by the factor analysis of items from a number of different scales (Ruth, 2006). The most popular instrument is the Mindful Attention Awareness Scale (MAAS) (Brown & Ryan, 2003).

Interestingly, while higher MAAS scores have been found to be associated with meditation experience in some studies (Brown & Ryan, 2003), other research using student samples found no significant correlation between MAAS scores and experience with meditative practices (assessed by self report) (MacKillop & Anderson, 2007). This finding which was also replicated by Thompson and Waltz (Thompson & Waltz, 2007) who suggest that this may be due to the fact that mindfulness during meditation may be a state like construct, while mindfulness in everyday life may be a trait like construct. While various rigorous reviews of clinical trials have found MBSR programmes to be effective for reducing stress (Chiesa & Serretti, 2009; Praissman, 2008), it is still unclear whether individuals with higher levels of mindfulness are also psychologically healthier. MBSR programs have also been associated with higher levels of optimism (Carson, Carson, Gil, & Baucom, 2004), but mindfulness and optimism have not tended to be researched together. Only one study where mindfulness levels were correlated positively with self reported health (Hansen, Lundh, Homman, & Wångby-Lundh, 2009) was identified as part of this research. Hansen, using a cross-sectional design, found that MAAS scores correlated with a five point one item measure of health, a very crude measure of health status. In a recent meta-analysis of 29 studies, Giluk explored the relationship between mindfulness, Big

Five personality and affect (Giluk, 2009). These authors found that mindfulness was strongly correlated with Neuroticism and negative affect, though its not obvious whether being mindful lowers neuroticism or whether neuroticism interferes with mindfulness.

4.4 Aims of the Research

As discussed above, optimism is well known as a predictor of health, both self reported and objectively assessed. Mindfulness, while a much newer construct, has also been associated with health in a number of studies. In addition, optimism has been associated with the placebo response in some research, while mindfulness has been proposed as a mediator of the relationship between explicit and implicit measures.

The primary aim of this study was to develop and test psychometric models that could be used to predict scores according to IRT and factor analysis criteria in the experimental part of the research. In order to do so, both factor analytic and item response models were fit to the optimism and mindfulness scales, as well as the health scales for Sample One.

A secondary aim of this part of thesis was both to replicate the optimism-health link reported in previous studies and to examine the relationships between health, mindfulness and optimism.

The major hypotheses of this part of the thesis were as follows:

- The RAND MOS would have 8 first order factors
- The MAAS would have one factor
- The LOT-R would have one factor.
- Optimism and mindfulness would be positively associated with health.
- Optimism and mindfulness would be positively associated with one another.

The response rate for Sample One (paper) was approximately 90% of those asked (N=392), while the response rate for Sample Two (Online) was 10% (N=1501).

4.5 Methods

The methods used for this part of the thesis were primarily psychometric. Cross-validation approaches (described in Chapter 3) were applied to both samples to increase generalisability of the models to the experimental portion of the research. This section describes the measures used for this part of the study, followed by a

description of the sample, and concludes with a description of the methods of analysis used in this study.

4.5.1 Participants

Table 4.1: Gender Breakdown of Participants by Collection Method

Gender	Online	Paper
Female	798	286
Male	309	104

As can be seen from Table 4.1, both samples were predominantly female, in line with the general gender balance of the University from which the samples were taken.

The majority of the first sample (N=390) were undergraduate students (N=304), and the mean age in the sample was 23.18 (SD=7.9). The majority of the second sample (N=1107) were undergraduate students (N=937) while the mean age was 22.3 (SD=6.8).

4.5.2 Measures

There were three measures used for this part of the research.

1. *RAND-MOS*: The RAND Medical Outcomes Survey produced the most widely used instrument of HRQoL (Health Related Quality of Life) worldwide (Hays, Sherbourne, & Mazel, 1993). The instrument was later revised and the number of response categories standardised across scales and renamed the SF-36. The older version was used for this research, as they are extremely similar and the newer version is under copyright and expensive to use, even for non-commercial research. The RAND-MOS has 36 questions, and is divided into 8 sub-scales, General Health (GH), Physical Functioning (PF), Role Limitations (RL), Emotional Role Limitations (RLE), Pain (PN), Energy (EN), Emotional Well Being (EMWB) and Social Functioning (SF). All sub-scales have shown acceptable reliability ($>.7$) in other studies in the literature (Lam, Chiu, & Lau, 2007; Ferreira, 2000). The instrument has 8 first order factors and two higher order factors (Hann & Reeves, 2008). The scale involves dichotomous, trichotomous and five and six point scales for various items, so all questions were transformed to a 100 point scale before analysis, where higher scores represent better functioning.
2. The *Mindful Attention Awareness Scale (MAAS)* (Brown & Ryan, 2003) is a 15 item scale which is scored on a six point scale from “almost always” to “almost never”. The scale uses questions which measure mindlessness. The summary score is produced from the mean of all individual scores. The scale has shown

adequate psychometric validity in many samples, with alpha ranging from 0.7 to 0.9 (Brown & Ryan, 2003; Ruth, 2006).

3. *Life Orientation Test, Revised*: The Life Orientation Test Revised (LOT-R) was developed and revised by Scheier and Carver (Scheier et al., 1994), and consists of 10 items. Three of the items load on pessimism, three on optimism and four are distractor items. The LOT-R has shown excellent psychometric validity, and is very commonly used as a measure of optimism/pessimism. The scale is scored on a 5 point scale, and the mean of all items after items 3,7 and 9 are reverse coded (and the distractor items removed) is taken to produce the overall score.

4.5.3 Sampling for this research

The participants in the paper collection method part completed the forms by hand between August and October 2009. The participants were sampled pseudo-randomly from all of the public areas (coffee shops, restaurants etc) of the campus.

Following this pen and paper approach to sampling, the survey was sent to a random selection of students via email on December 12th 2009, and data was collected and analysed from this point until the 24th of December 2009. Differences between the samples and the possible effects of these on the results obtained are discussed below.

4.6 Analysis

Analysis was carried out separately on the two samples, to allow for development of factor analysis and IRT models on the first sample and validation on the second. In addition, it could not be assumed that two samples collected in different ways would be comparable.

All missing data was assumed to be Missing Completely at Random (MCAR) (Little & Rubin, 1987), and thus a complete-cases analysis was carried out on all data.

The majority of the analytical procedures were as described in Chapter 3. In brief, Factor analytic and item response theory (not Rasch) models were fit to the data, and tested on either new samples (Sample One) or on other splits of the data (Sample Two). All regression results were obtained through the use of ten-fold cross-validation on the training set with a stepwise selection method, and the reported coefficients are from the test set, and thus are unbiased by the search process. Stepwise selection methods were used as they are reasonably standard within the field, and unlike other selection methods (lasso and ridge regression), these report a p-value, in line with APA standards. In general, all correlations reported are rank-correlations, to control for the non-normality of the input variables.

4.7 Results

4.7.1 Descriptive Statistics

In advance of the analysis, frequencies, means and ranges were calculated for the major variables of interest (General Health, Mindfulness, Optimism and Emotional Well Being) for the first sample. The results of this analysis can be seen in Table 4.2.

The mean level of mindfulness in the sample was higher than the mid-point of the scale, while optimism levels were at the mid-point point of the scale. The average level of self-reported health tended to be well above the mid-point for all of the subscales, which makes sense given the non-clinical sample involved in this research.

Table 4.2: Summary Statistics for Health Scales, Mindfulness and Optimism, Sample One

	Mean	SD	Median	Min	Max	Alpha
General Health	56.77	14.67	56.00	11.00	100.00	0.74
Physical Functioning (RAND)	86.19	22.26	95.00	0.00	100.00	0.93
Role Limitations (RAND)	85.22	26.68	100.00	0.00	100.00	0.75
Emotional Role Limitations	76.74	34.09	100.00	0.00	100.00	0.73
Energy Fatigue	56.25	16.97	60.00	0.00	100.00	0.72
Emotional Well Being	70.53	17.32	76.00	4.00	100.00	0.81
Social Functioning	80.81	20.76	87.50	0.00	100.00	0.61
Pain	81.84	19.55	90.00	0.00	100.00	0.77
Mindfulness (MAAS)	3.98	0.89	4.07	1.20	6.00	0.88
Optimism (LOT-R)	2.55	0.84	2.50	1.00	5.00	0.79

Table 4.3: Summary Statistics for Health Scales, Mindfulness and Optimism, Sample Two

Scale	Mean	SD	Median	Min	Max	Alpha
General Health	55.53	15.15	55.80	1.00	87.50	0.80
Physical Functioning (RAND)	90.74	17.50	95.00	0.00	100.00	0.93
Role Limitations (RAND)	80.94	31.70	100.00	0.00	100.00	0.69
Emotional Role Limitations	60.78	42.40	66.67	0.00	100.00	0.82
Emotional Well Being	65.37	19.51	68.00	4.00	100.00	0.85
Social Functioning	74.59	24.46	75.00	0.00	100.00	0.80
Pain	81.12	19.73	90.00	0.00	100.00	0.84
Energy Fatigue	49.74	20.27	50.00	0.00	100.00	0.79
Mindfulness (MAAS)	4.01	0.94	4.07	1.00	6.00	0.90
Optimism (LOT-R)	2.65	0.94	2.50	1.00	5.00	0.86

The summary statistics for the Sample Two are shown in Table 4.3. Emotional Role Limitations is lower in Sample Two. Note that zeros in some of the totals were due to some respondents only completing a subset of questions.

4.7.2 Psychometric Analyses

The primary aim of this chapter was to develop psychometric models for the mindfulness and optimism measures that could be used to supplement the data collected as part of the experimental research. Each scale, across both samples was taken as one unit, and factor analytic, item response methods, and structural equation models were used to develop and test combinations of items which seemed to provide useful explanatory power. Additionally, the use of multiple methods here allows for the actual impact of the different models to be assessed against one another, and those built on different sub-samples, to ensure that these models are reproducible.

This section covers the RAND-MOS data for Sample One (this was not repeated for Sample Two, as this instrument was not used in the experimental portion of the research). Next, the MAAS is investigated over both samples using all of these methods, followed by an examination of the LOT-R. The most successful models from this process were then tested against Sample One data in a back testing process.

4.7.3 RAND MOS

The parallel analysis criterion and the Kaiser criterion suggested eight factors. The acceleration factor and the optimal calibration index measures suggested two factors. These two measures may be picking up on the higher order factor structure of the items, as the RAND is typically modelled as having two higher order factors (physical and mental health).

Two and eight factor structures were examined for the RAND MOS and their performance assessed on unseen data (from sample 2) to determine which of these provides the best fit.

The results of the two factor solution (Appendix A, Table A.1, Page 202) were as follows: The first factor appears to contain all of the scales except for Physical Functioning, which loads on Factor 2. This factor can be best termed as General and Emotional Health.

The second factor maps exactly to the Physical Functioning Scale, and so retains that name. The non-normed fit index was equal to 0.693 and the RMSEA was equal to 0.095, with confidence intervals from 0.091 to 0.096.

This factor solution does not appear to be useful, as it has extremely low fit indices, and the breakdown of the factors is rather strange. If the factors had broken down in terms of Physical and Mental Health, then this would have made more sense. The factor loadings were invariant under a number of rotations (varimax, oblimin and promax), so it appears to be a real (if less than interpretable) factor structure.

Table 4.4: Factor Loadings Eight Factor Solution, RAND MOS, Sample One

	PhysFunc	SocEmWB	GenHealth	EmRoleLim	RoleLim	Fatigue	Pain	Energy	Communalites
RANDQ1	0.07	-0.01	0.54	0.08	0	-0.02	0.1	0.2	0.46
RANDQ2	0.01	-0.03	-0.07	0.06	0.16	-0.07	0.05	0.34	0.15
RANDQ3	0.38	-0.1	0.29	0.07	-0.08	-0.05	0.14	0.08	0.28
RANDQ4	0.74	0.03	0.04	0.03	0.01	0	0.02	-0.08	0.59
RANDQ5	0.71	0.06	0.05	-0.07	0.03	-0.01	0.05	-0.07	0.56
RANDQ6	0.69	-0.12	0.13	0.1	-0.08	-0.04	0.11	0.09	0.52
RANDQ7	0.89	0.04	-0.02	-0.02	-0.02	-0.05	0	-0.05	0.79
RANDQ8	0.75	-0.03	0.01	0	-0.08	0.05	0.1	-0.01	0.57
RANDQ9	0.81	-0.02	0.04	0	-0.01	0.05	0.01	0.05	0.67
RANDQ10	0.79	0	-0.03	0.02	0.03	-0.01	0	0.07	0.64
RANDQ11	0.87	0.03	-0.07	-0.04	0.05	0.01	-0.06	-0.01	0.76
RANDQ12	0.83	0	-0.09	0.01	0.11	0.04	-0.13	0	0.72
RANDQ13	0.01	-0.05	0.06	0.04	0.72	0	0.02	0.03	0.56
RANDQ14	0.01	-0.13	0.06	0.16	0.58	0.17	-0.05	0.01	0.45
RANDQ15	0.07	0.04	0	-0.06	0.66	-0.07	0.12	-0.03	0.52
RANDQ16	0	0.12	-0.01	0	0.6	0	0.1	0	0.46
RANDQ17	0.02	0.07	-0.01	0.69	0.08	-0.04	0	0.01	0.56
RANDQ18	-0.05	0.01	-0.02	0.78	0	0.06	-0.03	0.05	0.64
RANDQ19	0.04	-0.02	0.03	0.66	-0.01	0.05	-0.04	-0.06	0.43
RANDQ20	0.05	0.38	0.03	0.4	0.05	-0.07	0.27	-0.11	0.67
RANDQ21	0.01	-0.04	0.03	-0.01	-0.02	0.07	0.78	-0.02	0.63
RANDQ22	-0.02	0.01	-0.04	-0.02	0.13	0.01	0.8	0.03	0.71
RANDQ23	-0.06	0.32	0.11	0.11	0.03	0.16	0.01	0.31	0.46
RANDQ24	-0.03	0.52	-0.02	0.04	0	0.13	0.01	-0.12	0.36
RANDQ25	0.04	0.73	0.06	0.08	0.03	0.06	-0.02	-0.12	0.69
RANDQ26	-0.03	0.56	0.02	0.05	0.02	0.11	0.01	0.26	0.55
RANDQ27	-0.02	0.34	0.16	0.03	-0.02	0.21	0.13	0.37	0.61
RANDQ28	0.03	0.67	0.03	0.13	-0.05	0.1	0.02	-0.09	0.66
RANDQ29	0.05	0.12	-0.03	0	0.03	0.78	0.04	-0.06	0.73
RANDQ30	0.04	0.72	0.08	0.01	0.01	-0.01	-0.01	0.21	0.68
RANDQ31	-0.01	-0.05	0.04	0.04	-0.01	0.83	0.03	0.05	0.71
RANDQ32	0.07	0.39	0.04	0.18	0.06	0.01	0.17	-0.08	0.41
RANDQ33	0	-0.04	0.54	-0.04	0.05	0.19	0.04	-0.22	0.4
RANDQ34	-0.03	0.04	0.74	-0.02	0.05	-0.02	-0.03	-0.08	0.52
RANDQ35	0	0.14	0.29	-0.03	-0.01	0.13	-0.02	0.03	0.18
RANDQ36	0.03	-0.06	-0.79	-0.01	-0.02	0	0.02	-0.06	0.68

The factor correlations were quite low for this solution, at 0.27. This suggests that an orthogonal rotation might be more appropriate, but attempting this did not change any of the loadings. It seems that this factor structure is not theoretically or conceptually useful, which is an important feature of any proposed psychometric model.

Table 4.4 on Page 77 shows the loadings and has the named factors for the eight factor solution.

PhysFun: the first factor extracted maps exactly to the Physical Functioning scale, and retains that name.

SocEmWB: This scale maps to the Social Functioning, and the Emotional Well Being Scale. There are some items taken from the Energy/Fatigue scale, and as these are the positively worded items, this scale can probably best be termed as Social and Emotional Well Being.

GenHealth: These items map exactly to the General Health scale. Item 35 also loads on this scale, and as its loading was 0.29 while the cutoff was 0.30, it can be best characterised as part of that scale. Therefore, this factor can be best termed as General Health.

EmRoleLim: This scale maps to the Emotional Role Limitations and one item (20) from the Social Functioning scale (which asks about social events that have been missed due to health problems) and so this can probably best be termed as Emotional Role Limitations.

RoleLim: This scale maps exactly to the Role Limitations sub-scale, and so retains that name.

Fatigue: These items are the negative items from the Energy/Fatigue scale, and so this factor can probably best be termed as Fatigue.

Pain: These items map exactly to the Pain scale, and so retain that name.

Energy: The loadings on this factor are all below .4, which argues against its unproblematic interpretation. In addition, Q2 loads on this factor, when it is not typically associated with any factor. The other two items are the positively worded items from the Energy/Fatigue scales, and so this factor can best be termed Energy.

The non-normed fit index was equal to 0.901 and the RMSEA was equal to 0.054, with confidence intervals from 0.049 to 0.057.

This factor structure definitely makes sense, and the fit indices are acceptable, although the RMSEA is a little higher than would be wanted.

This factor structure seems appropriate for the data, and matches the number of factors proposed for this instrument. However, the interesting part is how it is different from the published sub-scales. To start, Social Functioning and Emotional Well Being have been merged. This makes sense, as the Social Functioning scale is quite small, and there may not have been enough data available to precisely segment them. However, it appears that Energy and Fatigue are perceived differently by this sub-population. The discrepancies here may have arisen from the age of the sample -

while the instrument was developed on a sample from the general population, the sample here was predominantly young, and such differences from the scale could represent this fact. Additionally, the sample was that of university students, which may also have affected the responses.

Table 4.5: Factor Correlations, Eight Factor Solution RAND MOS, Sample One

	PhysFunc	SocEmWB	GenHealth	EmRoleLim	RoleLim	Fatigue	Pain	Energy
PhysFunc	1.00	0.14	0.07	0.11	0.29	0.08	0.20	-0.04
SocEmWB	0.14	1.00	0.41	0.67	0.18	0.54	0.28	0.21
GenHealth	0.07	0.41	1.00	0.34	0.20	0.40	0.42	0.29
EmRoleLim	0.11	0.67	0.34	1.00	0.39	0.45	0.28	0.16
RoleLim	0.29	0.18	0.20	0.39	1.00	0.21	0.48	-0.01
Fatigue	0.08	0.54	0.40	0.45	0.21	1.00	0.31	0.15
Pain	0.20	0.28	0.42	0.28	0.48	0.31	1.00	0.12
Energy	-0.04	0.21	0.29	0.16	-0.01	0.15	0.12	1.00

The factor correlations are shown in Table 4.5 and were moderate ($r = 0.1 - 0.5$), and in line with expectations.

Table 4.6: SEM Comparison for RAND MOS Factor Solutions, Sample One

base	comparison	ep	minus2LL	df	AIC	diffLL	diffdf	p
RAND2Samp1		68	221132.03	598.00	130190.69			
RAND2Samp1	RAND8Samp1	75	248722.59	591.00	157795.25	27590.56	-7.00	

As can be seen from Table 4.6, the 8 factor solution appears to fit better (lower AIC), so on the basis of this analysis, this is the solution which should be retained.

Table 4.7: Model Comparison for RAND MOS data using Models from Sample 1 on Sample 2

base	comparison	ep	minus2LL	df	AIC	diffLL	diffdf	p
RAND2Samp2		68	167535.98	598.00	98644.14			
RAND2Samp2	RAND8Samp2	75	188653.78	591.00	118246.73	21117.80	-7.00	

As Table 4.7 shows, the eight factor model provided the best fit to this unseen data (Split A of the second Sample), which is similar to the data on which the model was created. This supports the hypothesis proposed in the introduction, and is in line with previous research.

4.8 Mindfulness Attention Awareness Scale

The next step in the analyses was to examine the MAAS using factor analytic and IRT methods.

Table 4.8 shows the item content for the MAAS.

4.8.1 Factor Analyses, Sample One

For the MAAS, parallel analysis, the MAP, VSS and Kaisers rule methods suggested a one factor solution, while the acceleration factor and optimal coordinates index

Table 4.8: MAAS Item Content

Item	Content
MAASQ1	I could be experiencing some emotion and not be conscious of it until some time later
MAASQ2	I break or spill things because of carelessness, not paying attention, or thinking of something else.
MAASQ3	I find it difficult to stay focused on what's happening in the present.
MAASQ4	I tend to walk quickly to get where I'm going without paying attention to what I experience along the way
MAASQ5	I tend not to notice feelings of physical tension or discomfort until they really grab my attention
MAASQ6	I forget a person's name almost as soon as I've been told it for the first time.
MAASQ7	It seems I am "running on automatic," without much awareness of what I'm doing.
MAASQ8	I rush through activities without being really attentive to them
MAASQ9	I get so focused on the goal I want to achieve that I lose touch with what I'm doing right now to get there.
MAASQ10	I do jobs or tasks automatically, without being aware of what I'm doing.
MAASQ11	I find myself listening to someone with one ear, doing something else at the same time.
MAASQ12	I drive places on 'automatic pilot' and then wonder why I went there.
MAASQ13	I find myself preoccupied with the future or the past
MAASQ14	I find myself doing things without paying attention
MAASQ15	I snack without being aware that I'm eating

methods suggested a three factor solution. Therefore, one and three factor solutions were extracted and the results interpreted, as shown below.

Table 4.9: Factor Loadings, One Factor Solution, MAAS, Sample One

	Mindfulness	Communalites
MAASQ1	0.48	0.24
MAASQ2	0.44	0.19
MAASQ3	0.59	0.35
MAASQ4	0.56	0.31
MAASQ5	0.51	0.26
MAASQ6	0.36	0.13
MAASQ7	0.73	0.54
MAASQ8	0.78	0.6
MAASQ9	0.62	0.39
MAASQ10	0.72	0.52
MAASQ11	0.54	0.29
MAASQ12	0.57	0.32
MAASQ13	0.58	0.33
MAASQ14	0.76	0.58
MAASQ15	0.45	0.21

The results of the one factor solution for the MAAS are shown in Table 4.9.

The non-normed fit index was equal to 0.874 and the RMSEA was equal to 0.078, with confidence intervals from 0.067 to 0.087.

The next factor solution to be examined was the three factor solution (loadings shown in Chapter A, Table A.3, Page 203)

PA1: "Q8" "Q10" "Q11" "Q12" "Q13" "Q14" "Q15" All of these questions relate to lack of awareness, and so this factor can best be termed this.

PA3: "Q4" "Q5" "Q6" "Q7" "Q8" "Q9" These mostly relate to sensations of physical unawareness, and so this factor can best be termed physical unawareness.

PA2: "Q1" "Q2" "Q3" "Q14". This factor can perhaps best be termed as lack of

present attention.

While the amount of variance explained increased with the number of factors, this model does not seem to be particularly useful, in that it does not shed new light on the construct.

The non-normed fit index was equal to 0.937 and the RMSEA was equal to 0.056, with confidence intervals from 0.041 to 0.067.

Table 4.10: Comparison of One and Three Factor Models, Sample One

base	comparison	ep	minus2LL	df	AIC	diffLL	diffdf	p
MAAS1		30	51805.39	90.00	44016.61			
MAAS1	MAAS3	32	8983.52	88.00	1198.75	-42821.87	-2.00	

Factor solutions for one and three factors were extracted, and the results were subjected to CFA.

From Table 4.10 it can be seen that the best model is the one factor model, which is in line with previous research. The factor structure is not reported here as all factors loaded on the first factor. This factor explained 35% of the variance in the sample, which is low.

Table 4.11: Comparison of Sample One MAAS Factor Models on a subset of Sample Two Data (Split A)

base	comparison	ep	minus2LL	df	AIC	diffLL	diffdf	p
MAAS1Samp2		30	6361.20	90.00	64.48			
MAAS1Samp2	MAAS3Fit2	32	7355.74	88.00	1063.03	994.55	-2.00	

Table 4.11 demonstrates that the MAAS 1 factor model provided the best fit to the subsample of data ($N = 313$) used to test the model.

4.8.2 Factor Analyses, Sample Two

The parallel analysis procedure for Split B suggested that this sample of the responses to the MAAS has five factors, while the MAP criterion suggests that it has only one. Following our previous approach, each of these factor solutions will be examined and interpreted before a CFA is applied on the remainder of the dataset.

For split C, the same procedure was carried out, and the various methods both suggested five and one factors, respectively.

Table 4.12 shows the one factor loadings averaged over Splits B and C. The solution only explained 36% of the variance in the item loadings, which is in line with the first sample, though much lower than the original published research in which the MAAS was developed.

Table 4.12: One Factor Solution Averaged over Splits B and C, MAAS

	Mindfulness	Communalities
MAASQ1	0.53	0.28
MAASQ2	0.51	0.26
MAASQ3	0.71	0.51
MAASQ4	0.55	0.3
MAASQ5	0.53	0.28
MAASQ6	0.29	0.08
MAASQ7	0.77	0.59
MAASQ8	0.79	0.62
MAASQ9	0.67	0.45
MAASQ10	0.74	0.55
MAASQ11	0.59	0.35
MAASQ12	0.59	0.35
MAASQ13	0.59	0.35
MAASQ14	0.8	0.64
MAASQ15	0.46	0.21

This five factor solution for Split B (Appendix A, Table A.5, page 204) explained 54% of the variance in the sample.

PA1: “Q5”, “Q6”, “Q7”, “Q8”, “Q9”, “Q10”. This factor has come through in most of the previous solutions, and can again be termed distractability.

PA2: “Q1”, “Q2”, “Q3”. Again, these items have clustered together previously, and this factor is again termed lack of present awareness.

PA3: “Q4”, “Q5”. This factor is again termed lack of somatic awareness.

PA4: “Q13”, “Q14”. This factor can best be termed as lack of attention.

PA5: “Q10”, “Q11”, “Q12”, “Q14”, “Q15”. This factor again can be termed distractability.

Although there is significantly more variance explained by this solution, again it does not lend any more conceptual clarity to the instrument, suggesting that it is not worth the extra parameters estimated. Note that Q1-3 clustered together in previous multi-factor solutions, so there may be a different construct underlying these questions, but this is not clearly replicable across the entire sample.

The five factor solution for split C (Appendix A, Table A.6, Page 204) broke down as follows:

PA1: “Q7” “Q8” “Q9” “Q10” “Q11” “Q12” “Q14” All of the items in this factor relate to a lack of attention to the present, and it can probably be best termed lack of present focus.

PA2: “Q9” “Q13” Q9 loads on both PA1 and PA2, and as PA2 has really only Q13 loading to any major extent on it, no interpretation of this factor was performed. It

was named lack of present awareness.

PA4: "Q2" "Q3" "Q14" This factor can probably best be termed lack of attention.

PA5: "Q1" "Q4" "Q5" Most of these items relate to lack of bodily attention, and this can probably best be termed lack of somatic awareness.

PA3: "Q4" "Q6" "Q7" "Q8" Note that MAASQ4 loads slightly on both PA3 and PA5, and is not considered in the interpretation here. These items can probably best be termed lack of awareness.

Table 4.13: Comparison of Factor Structures for MAAS 2B Solutions, Tested on Split B

base	comparison	ep	minus2LL	df	AIC	diffLL	diffdf	p
MAAS12b		30	52617.86	90.00	45289.97			
MAAS12b	MAAS52b	32	8651.98	88.00	1328.08	-43965.89	-2.00	

As can be seen from Table 4.13, the one factor solution again performs best, further increasing our confidence in its adequacy.

Table 4.14: Comparison of Factor Structures for MAAS 2C Solutions, Tested on Split C

base	comparison	ep	minus2LL	df	AIC	diffLL	diffdf	p
MAAS12c		30	7524.73	90.00	166.35			
MAAS12c	MAAS52c	38	8345.87	82.00	1003.48	821.13	-8.00	

As can be seen from Table 4.14, the one factor model was again superior to the five factor model in Split C.

4.8.3 IRT Analyses - MAAS

4.8.3.1 Sample One

The MAAS was then considered from a IRT perspective. Firstly, the instrument was examined using Mokken analysis to check if it could be considered one scale, and whether or not there were violation of monotonicity.

The Mokken analysis suggests that two items should be dropped from the scale, items 2 and 6² (Van Der Ark, 2007). This leaves a thirteen item scale for further analysis. Note that items 2 and 6 had extremely low communalities in Table 4.12, which reinforces the decision to remove these items from the scale.

There were no violations of the monotonicity assumption for the reduced scale. The item coefficients (ItemH) are quite low, many of them hang around 0.30, which is the minimum allowed.

²because they do not appear to be part of the scale

The next stage of analysis for the Mindfulness scale was to fit one and two parameter Graded Response Models to the data.

Table 4.15: Coefficient Estimates for MAAS One Parameter Graded Response Model, Sample One

	$\beta^1(se)$	$\beta^2(se)$	$\beta^3(se)$	$\beta^4(se)$	$\beta^5(se)$	α
MAASQ1	-2.579 (0.21)	-1.795 (0.24)	-0.733 (0.37)	-0.176 (0.19)	0.806 (0.28)	1.525 (0.07)
MAASQ3	-2.482 (0.2)	-1.573 (0.24)	-0.423 (0.4)	0.293 (0.83)	1.52 (0.07)	1.525 (0.21)
MAASQ4	-1.72 (0.17)	-0.649 (0.24)	0.161 (0.4)	0.74 (0.07)	1.851 (0.25)	1.525 (0.18)
MAASQ5	-2.477 (0.17)	-1.293 (0.26)	-0.56 (0.07)	0.023 (0.23)	1.191 (1.33)	1.525 (0.16)
MAASQ7	-2.674 (0.24)	-1.514 (0.07)	-0.544 (0.25)	0.334 (0.21)	1.536 (1.31)	1.525 (0.18)
MAASQ8	-3.012 (0.07)	-1.677 (0.2)	-0.612 (0.2)	0.314 (0.2)	1.802 (1.31)	1.525 (0.31)
MAASQ9	-2.606 (0.2)	-1.569 (1.32)	-0.654 (0.19)	0.104 (0.2)	1.386 (1.31)	1.525 (0.07)
MAASQ10	-2.814 (0.19)	-1.636 (1.31)	-0.479 (0.6)	0.336 (0.22)	1.674 (0.07)	1.525 (0.18)
MAASQ11	-2.353 (0.17)	-1.15 (1.31)	-0.022 (0.61)	0.737 (0.07)	2.178 (0.14)	1.525 (0.14)
MAASQ12	-2.918 (0.17)	-1.8 (1.31)	-0.955 (0.07)	-0.517 (0.19)	0.286 (0.16)	1.525 (0.17)
MAASQ13	-1.624 (0.2)	-0.609 (0.07)	0.316 (0.21)	0.93 (0.27)	1.894 (1.57)	1.525 (0.37)
MAASQ14	-2.609 (0.07)	-1.43 (0.22)	-0.295 (0.19)	0.506 (0.27)	1.902 (1.55)	1.525 (1.12)
MAASQ15	-2.186 (0.14)	-1.29 (0.14)	-0.765 (0.17)	-0.323 (0.27)	0.468 (1.55)	1.525 (0.07)

Table 4.15 shows the estimated ability thresholds and discrimination parameter for the one parameter Graded Response Model on the MAAS. The discrimination parameter is moderate, as are the ability estimates, suggesting that this scale may not be suitable for respondents particularly high in mindfulness.

Next, a two parameter Graded Response Model was examined for the same scale.

The two parameter Graded Response Model (Appendix A, Table A.7, Page 205) had the following features. Q14 has the highest discriminatory power, and that Q11 has the highest ability threshold, while Q1 has the lowest. Q11 refers to listening to others while engaging in other tasks, and its ability estimates suggest that it is a good question for pinpointing the abilities of respondents high on the construct of mindfulness.

The item information curves for each item in one and two parameter models can be seen in Figure 4.1. Note that items 7, 8, 9, 10 appear to convey far more information over the participant ability levels. However, as seen below, this model was not supported on unseen data.

Table 4.16: Performance of MAAS One and Two Parameter GRM's on Unseen Data (Sample Two, Split A)

	ErrorApproximation	Correlation
One Parameter GRM	9.44	1.00
Two Parameter GRM	12.62	1.00

As can be seen from Table 4.16, the one parameter GRM provided a better fit to the unseen data, in contrast to the likelihood-based method. This suggests again that a simpler model for the MAAS, in line with previous research, appears to be a better model for the items.

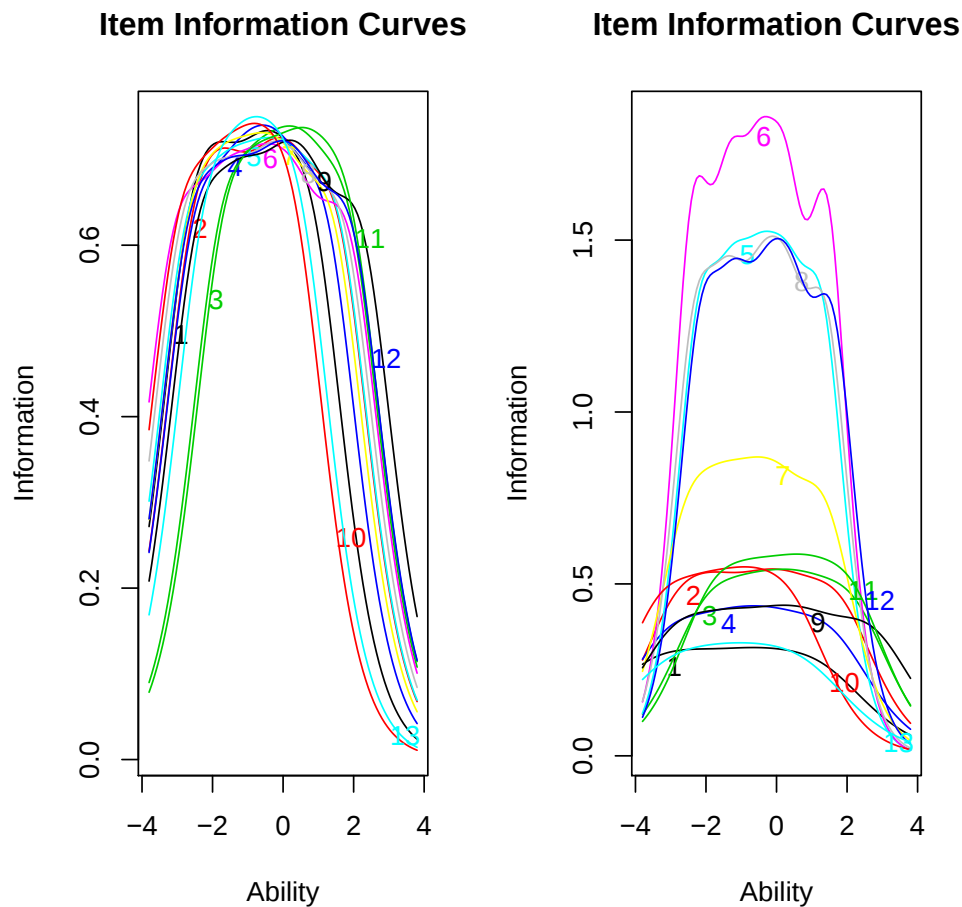


Figure 4.1: Item Information Curves for MAAS One Parameter GRM (left) and Two Parameter GRM (right) (Sample One)

4.8.3.2 Sample Two

First, the assumptions underlying item response theory modelling are checked.

Questions 5 and 6 failed the item ordering assumptions for Split B and so are removed from the scale before further analysis. Note that Q6 was problematic in the factor analytical solutions too, although Q5 was in line with retained items in the factor analysis. The reduced scale had no violations of the monotonicity assumption.

For Split C, items 5 and 11 fail the item ordering check, and so are removed. The reduced scale had no failures of the monotonicity assumption, so modelling continues with the reduced scale.

Next, one and two parameter graded response models were fit to Split B.

It can be seen from Table 4.17 that there were no problems with the parameter estimates for this model.

Table 4.17: Coefficient Estimates for MAAS One Parameter Graded Response Model, Split B

	$\beta^1(se)$	$\beta^2(se)$	$\beta^3(se)$	$\beta^4(se)$	$\beta^5(se)$	α
MAASQ1	-3.084 (0.28)	-1.803 (0.34)	-0.513 (0.94)	0.026 (0.16)	0.961 (0.27)	1.735 (0.07)
MAASQ2	-2.872 (0.27)	-1.832 (0.34)	-0.99 (0.92)	-0.437 (0.57)	0.65 (0.07)	1.735 (0.18)
MAASQ3	-2.657 (0.26)	-1.427 (0.34)	-0.429 (0.92)	0.246 (0.07)	1.53 (0.25)	1.735 (0.17)
MAASQ4	-1.814 (0.27)	-0.819 (0.34)	0 (0.07)	0.581 (0.23)	1.624 (1.15)	1.735 (0.16)
MAASQ7	-2.515 (0.38)	-1.479 (0.07)	-0.505 (0.2)	0.227 (0.22)	1.262 (1.16)	1.735 (0.16)
MAASQ8	-2.615 (0.07)	-1.584 (0.14)	-0.532 (0.17)	0.18 (0.21)	1.409 (1.16)	1.735 (0.33)
MAASQ9	-2.337 (0.25)	-1.487 (0.88)	-0.7 (0.15)	-0.016 (0.21)	1.161 (1.16)	1.735 (0.07)
MAASQ10	-2.812 (0.25)	-1.422 (0.87)	-0.506 (0.56)	0.246 (0.22)	1.369 (0.07)	1.735 (0.15)
MAASQ11	-2.263 (0.24)	-0.936 (0.87)	-0.092 (0.57)	0.529 (0.07)	1.606 (0.11)	1.735 (0.12)
MAASQ12	-2.937 (0.23)	-1.944 (0.87)	-1.246 (0.07)	-0.793 (0.17)	-0.091 (0.13)	1.735 (0.15)
MAASQ13	-1.247 (0.23)	-0.375 (0.07)	0.392 (0.18)	1.003 (0.27)	1.872 (0.94)	1.735 (0.48)
MAASQ14	-2.394 (0.07)	-1.212 (0.19)	-0.261 (0.17)	0.356 (0.27)	1.606 (0.95)	1.735 (1.5)
MAASQ15	-1.911 (0.21)	-1.163 (0.17)	-0.529 (0.15)	-0.13 (0.27)	0.656 (0.95)	1.735 (0.07)

The two parameter GRM model (not shown) had no problems with the coefficient estimates for this solution, and the estimates show the usual trade off between discrimination and ability estimate parameters.

The item information curves for the Split B data (Figure 4.2) show a similar breakdown to those from Sample One. This highlights the benefits of fitting both these models, as if the two parameter model were true then approximately similar performance could be obtained from a much smaller scale. Unfortunately for this theory, the more complex model was not supported across any of the splits.

Next, we examine the performance of these models on unseen data.

Table 4.18: Performance of One and Two Parameter Graded Response Models for MAAS on Unseen Data (Splits A and C)

	ErrorApproximation	Correlation
One Parameter	5.02	1.00
Two Parameter	7.03	1.00

As can be seen from Table 4.18, the one parameter model provided the best fit to the unseen data.

Next, one and two parameter Graded Response Models are fit to Split C.

Table 4.19: Coefficient Estimates for MAAS One Parameter Graded Response Model, Split C

	$\beta^1(se)$	$\beta^2(se)$	$\beta^3(se)$	$\beta^4(se)$	$\beta^5(se)$	α
MAASQ1	-2.59 (0.23)	-1.61 (0.08)	-0.476 (1.18)	0.077 (0.71)	1.036 (0.25)	1.668 (0.73)
MAASQ2	-2.67 (0.22)	-1.759 (0.21)	-0.805 (0.08)	-0.148 (0.71)	0.719 (0.25)	1.668 (0.73)
MAASQ3	-2.502 (0.2)	-1.226 (0.31)	-0.304 (0.25)	0.479 (0.08)	1.727 (0.27)	1.668 (0.73)
MAASQ4	-1.81 (0.2)	-0.797 (0.3)	0.039 (0.22)	0.752 (0.25)	1.809 (0.08)	1.668 (0.73)
MAASQ7	-2.853 (0.37)	-1.583 (0.3)	-0.519 (1.21)	0.249 (0.23)	1.453 (0.28)	1.668 (0.08)
MAASQ8	-3.032 (0.08)	-1.619 (0.31)	-0.523 (1.19)	0.271 (0.22)	1.8 (0.56)	1.668 (0.23)
MAASQ9	-2.796 (0.24)	-1.553 (0.08)	-0.701 (1.19)	0.204 (0.22)	1.621 (0.55)	1.668 (0.2)
MAASQ10	-3 (0.23)	-1.664 (0.16)	-0.651 (0.08)	0.185 (1.05)	1.538 (0.55)	1.668 (0.18)
MAASQ12	-2.933 (0.22)	-1.962 (1.19)	-1.242 (0.28)	-0.662 (0.08)	0.047 (0.55)	1.668 (0.33)
MAASQ13	-1.264 (0.22)	-0.278 (1.18)	0.552 (0.23)	1.127 (0.28)	2.048 (0.08)	1.668 (0.33)
MAASQ14	-2.721 (0.22)	-1.515 (1.18)	-0.412 (0.21)	0.514 (0.27)	1.79 (0.12)	1.668 (0.08)

As can be seen from Table 4.19, the coefficient estimates appear reasonable. The

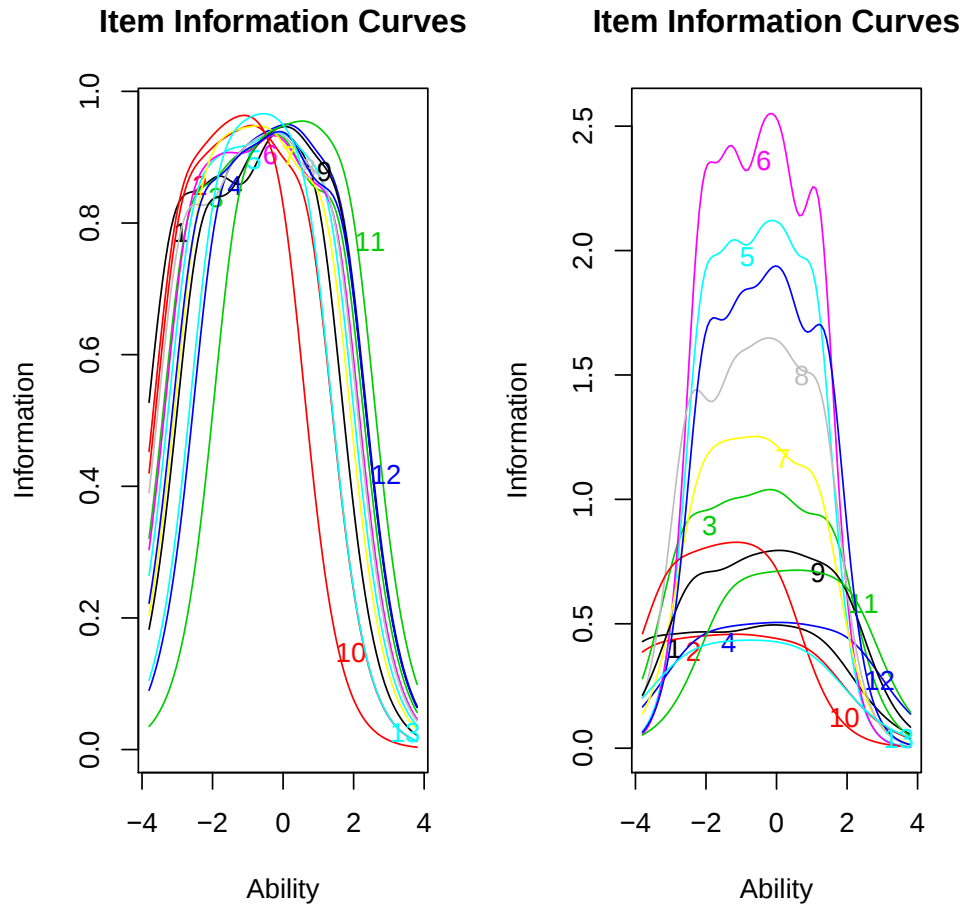


Figure 4.2: Item Information Curves for MAAS One Parameter GRM (left) and Two Parameter GRM (right) (Sample Two, Split B)

discrimination parameter is relatively low, suggesting that this scale is good for all levels of abilities, even though the highest estimated difficulty parameter is only 2.05, for Q13 which is “I often find myself occupied with the future or the past”, which is a relatively concise summary of the entire construct of mindlessness.

Next, a two parameter model was fit to this data.

The two parameter model (Appendix A, Page 206, Table A.9), is not that much different from the one parameter model. Of interest is that Q13 remains the most difficult question, but its discrimination parameter has come down, suggesting that it behaves similarly for participants of all ability levels. Q8 has the highest discrimination parameter of all the items and is “I rush through activities without being really attentive to them” and it appears that this question is the best at discrimination between those higher and lower on the construct of mindfulness.

Again, the item information curves for the two parameter model (Figure 4.3) show that items 3, 5, 6, 7, 8 and 11 appear to convey much more information across the

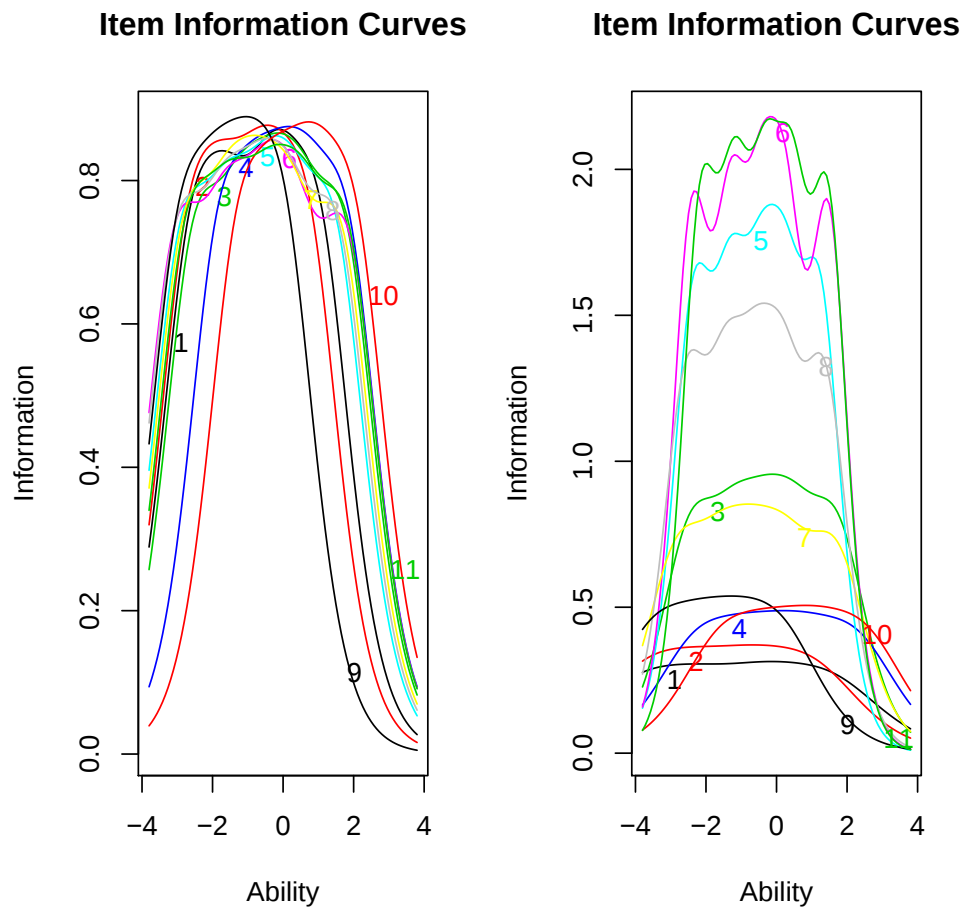


Figure 4.3: Item Information Curves for MAAS One Parameter GRM (left) and Two Parameter GRM (right) (Sample Two, Split C)

range of the test, suggesting that a reduced scale could convey more information with fewer items. However, the out of sample tests did not support the additional parameters for this model.

The final step in the analysis of the MAAS scale is to test the performance of the models on unseen data.

Table 4.20: Performance of MAAS One and Two Parameter Graded Response Models on Unseen Data (Splits A and B)

	ErrorApproximation	Correlation
One Parameter	13.59	1.00
Two Parameter	15.00	1.00

As can be seen from Table 4.20, the one parameter model provided a better fit to the unseen data (though neither model was particularly good). To finalise the examination of the MAAS, in all cases a one factor FA model and one parameter GRM provided the best fit to the data. Additionally, Q6 was removed by all of the

IRT models, suggesting that the responses to this question were not consistent with the responses to all of the other questions, and which questions this item's inclusion in the instrument (at least for this particular population).

4.8.4 Back testing

Finally, a back testing procedure was applied, where the best fitting models from the second sample were refit on data from the first sample. Both splits showed that a one factor solution and a one parameter Graded Response Model provided the best performance on unseen data.

Table 4.21: Full and Reduced Models for MAAS Tested on Sample One Data

base	comparison	ep	minus2LL	df	AIC	diffLL	diffdf	p
MAASBackTestFull		30	51358.59	90.00	43737.50			
MAASBackTestFull	MAASBackTestReduced1	26	27943.73	65.00	21597.15	-23414.86	-25.00	
MAASBackTestFull	MAASBackTestReduced2	22	36970.61	44.00	31664.97	-14387.98	-46.00	

As can be seen from Table 4.21, the reduced model without questions 5, 6 and 15 provided the best fit to the unseen data.

The next step in the back testing procedure was to test the IRT models on Sample One.

Table 4.22: Backtesting of MAAS IRT Models on Sample One Data

	ErrorApproximation	Correlation
Split B	7.00	1.00
Split C	2.19	1.00

It can be seen that the model developed from Split C (without items 5,6 and 15) provides the best fit to the unseen data, and this model will be used to generate scores for the experimental participants.

4.9 Life Orientation Test, Revised

The next step in the analysis was to investigate the structure of the LOT-R with respect to factor analysis and IRT methods.

Table 4.23 shows the content of the items for the LOT-R.

4.9.1 Factor Analyses, Sample One

Parallel Analysis indicated that two factors should be extracted, while the MAP criterion suggested one. Therefore, both one and two factor solutions were extracted from the matrix and their results examined for adequacy and interpretability.

Table 4.23: Life Orientation Test, Revised Item Content

Item	Question
LOTRQ1	In uncertain times, I usually expect the best
LOTRQ2	Its easy for me to relax
LOTRQ3	If something can go wrong for me, it will
LOTRQ4	I'm always optimistic about the future
LOTRQ5	I enjoy my friends a lot
LOTRQ6	Its important for me to keep busy
LOTRQ7	I hardly ever expect things to go my way
LOTRQ8	I don't get upset too easily
LOTRQ9	I rarely count on good things happening to me
LOTRQ10	Overall, I expect more good things to happen to me than bad

Table 4.24: Factor Loadings, One Factor Solution, LOT-R, Sample One

	Optimism	Communalites
LOTRQ1	0.49	0.24
LOTRQ3	0.6	0.35
LOTRQ4	0.61	0.37
LOTRQ7	0.77	0.59
LOTRQ9	0.61	0.37
LOTRQ10	0.65	0.43

Table 4.24 shows the loadings for the one factor solution. The communalities are relatively low, which is a sign that perhaps this solution is not optimal. As can be seen below, the two factor solution increases the communalities, but at the cost of potentially unwarranted complexity.

The non-normed fit index was equal to 0.853 and the RMSEA was equal to 0.124, with confidence intervals from 0.094 to 0.154. This solution does not seem optimal, as the RMSEA is well outside the recommended bounds, and the NNFI is quite low.

Table 4.25: Factor Loadings, Two Factor Solution, LOT-R, Sample One

	Pessimism	Optimism	Communalites
LOTRQ1	-0.07	0.71	0.46
LOTRQ3	0.8	-0.11	0.56
LOTRQ4	0.17	0.56	0.44
LOTRQ7	0.65	0.2	0.61
LOTRQ9	0.51	0.16	0.38
LOTRQ10	0.3	0.44	0.43

The two factor solution (Table 4.25) broke down as follows: PA2: "Q1", "Q4", "Q10". These items are all the positively framed items, and so this factor can best be termed as Optimism.

PA1: "Q3", "Q7", "Q9", "Q10" This factor consisted of the pessimism items, and so can best be termed Pessimism. Note that Item 10 has extremely poor loadings on both factors, which is surprising given that it can often be taken as an indicator for

the entire construct. Additionally, Q7 and Q10 have a small cross-loading on the optimism factor, and the communalities are quite low, suggesting that this model is not particularly well-fitting for this sample.

Table 4.26: Comparison of CFA for the LOT-R Sample One Models on Sample One

	base	comparison	ep	minus2LL	df	AIC	diffLL	diffdf	p
1	LOTR1		12	2506.12	9.00	40.33			
2	LOTR1	LOTR2	13	2878.10	8.00	414.31	371.98	-1.00	

As can be seen from Table 4.26, the one factor solution provided the best fit to the data. Therefore, this solution will be tested on the second sample.

Table 4.27: Comparison of CFA Results for LOTR Sample One, One and Two Factor Models on a Subset of Sample Two (Split A)

	base	comparison	ep	minus2LL	df	AIC	diffLL	diffdf	p
1	LOTR1		12	14395.44	9.00	12513.64			
2	LOTR1	LOTR2Model2	13	2376.67	8.00	496.87	-12018.77	-1.00	

Table 4.27, shows that the one factor model provides the best fit to the subsample of data used to examine the models performance on new data.

4.9.2 Factor Analyses, Sample Two

Again, the parallel analysis criterion suggests two factors, while the MAP criterion suggests one, so both solutions will be examined and interpreted for both splits, as the results were broadly similar.

Table 4.28: Average One Factor Solution for LOT-R Across Splits B and C

	Optimism	Communalities
LOTRQ1	0.55	0.3
LOTRQ3	0.73	0.53
LOTRQ4	0.64	0.42
LOTRQ7	0.78	0.61
LOTRQ9	0.74	0.54
LOTRQ10	0.77	0.6

Table 4.28 show the averaged coefficients across both splits. Note that the communalities are, on average, much higher than those observed in the first sample. This may be due to the larger sample size or potentially due to the pseudo-random sampling method used in the pen and paper sample. The two solutions are somewhat different. The communalities for Q1 are much higher in Split C than in Split B, while those for Q4 are much lower in Split C than in Split B. To some extent this probably represents sampling error, but it is quite strange that the two negatively worded items should show much of the variance across samples. The averaging produces a solution which appears much more stable.

Table 4.29: Average Two Factor Solution for LOT-R Across Splits B and C

	Optimism	Pessimism	Communalities
LOTRQ1	-0.08	0.75	0.49
LOTRQ3	0.8	-0.03	0.62
LOTRQ4	0.05	0.7	0.54
LOTRQ7	0.86	-0.02	0.72
LOTRQ9	0.64	0.13	0.56
LOTRQ10	0.33	0.51	0.6

Table 4.29 shows the averaged solution across Splits B and C. In general, though the major outline of the structure is the same, Split B's solution is much cleaner than that of Split C. For example, in Split C it is unclear whether or not Q10 belongs to the first or second factor. Therefore, it would seem advisable to prefer the structure from Split B, but this will be tested.

Table 4.30: Comparison of LOT-R Split B Factor Solutions, Tested on Split C

base	comparison	ep	minus2LL	df	AIC	diffLL	diffdf	p
LOTR1b		12	2137.26	9.00	44.53			
LOTR1b	LOTR2b	12	2739.52	9.00	646.80	602.27	0.00	0.00

Table 4.31: Comparison of LOT-R Split C Factor Solutions, Tested on Split B

base	comparison	ep	minus2LL	df	AIC	diffLL	diffdf	p
LOTR1c		12	2423.95	9.00	80.95			
LOTR1c	LOTR2c	12	2641.70	9.00	298.70	217.75	0.00	0.00

Tables 4.30 and 4.31 show the performance of the one and two factor solutions on the respective splits. In both solutions, the one factor solution performed better.

4.9.3 IRT Analyses, LOT-R

4.9.3.1 Sample One

Firstly, the scale analysis was conducted to determine which items fit best together.

All of the items meet the assumptions of a uni-dimensional scale (reinforcing the superiority of a one factor model seen in the earlier analyses). Next, the item orderings were examined.

Q1 was removed from the scale in order to meet the assumptions of the model (the IIO assumption). Note that this item had quite a low communality in the FA analyses also. There were no violations of monotonicity in the sample.

After this process of model checking, a five item scale remained for further analysis.

The next stage in the analysis of the LOT-R was the fitting of one and two parameter Graded Response Models.

An analysis of the one parameter GRM for the LOT-R showed that there were no violations of item ordering, and that LOTR4 is the hardest item to endorse, while LOTR3 appears to be the easiest.

The next step was to assess which of the models provided the best fit to the data using a likelihood ratio test.

The two parameter model was significantly better than the one parameter model ($p = 0.001$), but the AIC suggested that the one parameter model provided a better overall fit to the data.

Table 4.32: Coefficient Estimates for LOTR Two Parameter Graded Response Model

	$\beta^1(se)$	$\beta^2(se)$	$\beta^3(se)$	$\beta^4(se)$	α
LOTRQ3	-1.316 (0.14)	-0.284 (0.14)	0.603 (0.11)	1.711 (0.13)	1.759 (0.11)
LOTRQ4	-0.994 (0.11)	0.501 (0.11)	1.059 (0.1)	2.962 (0.08)	1.339 (0.17)
LOTRQ7	-1.084 (0.16)	-0.029 (0.45)	0.685 (0.25)	1.66 (0.27)	2.704 (1.01)
LOTRQ9	-1.061 (1.22)	0.085 (4.57)	0.849 (4.11)	1.93 (1.88)	1.726 (10.29)
LOTRQ10	-0.509 (0.19)	0.729 (0.16)	1.595 (0.33)	2.872 (0.19)	1.613 (0.19)

The estimates for the two parameter model can be seen in Table 4.32. Note that there are problematic estimates for Q4 at the third response point, suggesting that this model is not particularly useful. Again note how the FA analyses showed some problems with these questions, increasing our confidence in this analysis (for this particular sample, at least).

Finally, the performance of the one and two parameter IRT models for the LOT-R were assessed.

Table 4.33: Performance of LOTR One and Two Parameter GRM's on Unseen Data (Sample Two, Split A)

	ErrorApproximation	Correlation
One Parameter GRM	5.28	1.00
Two Parameter GRM	6.29	1.00

As can be seen from Table 4.33, the one parameter model provided a better fit to the unseen data.

4.9.3.2 Sample Two

Next, a similar procedure was followed for the second sample.

The LOT-R for Split B had no problems with either item ordering or monotonicity, suggesting that some of the problems observed earlier may have been due to variability from the particular sample.

Next, one and two parameter Graded Response Models were fit to the data.

Table 4.34: Coefficient Estimates for One Parameter Graded Response Model, LOT-R, Split B

	$\beta^1(se)$	$\beta^2(se)$	$\beta^3(se)$	$\beta^4(se)$	α
LOTRQ1	-1.345 (0.13)	-0.165 (0.11)	0.645 (0.24)	1.576 (1.64)	1.895 (0.1)
LOTRQ3	-1.278 (0.1)	-0.231 (0.11)	0.504 (0.26)	1.712 (0.1)	1.895 (0.11)
LOTRQ4	-0.985 (0.18)	0.262 (0.18)	0.856 (0.1)	1.885 (0.12)	1.895 (0.09)
LOTRQ7	-1.252 (1.2)	-0.165 (0.1)	0.719 (0.12)	1.644 (0.15)	1.895 (0.2)
LOTRQ9	-1.223 (0.1)	-0.109 (0.11)	0.69 (1.67)	1.653 (1.18)	1.895 (1.74)
LOTRQ10	-0.934 (0.13)	0.272 (0.24)	1.279 (1.64)	2.334 (1.18)	1.895 (0.1)

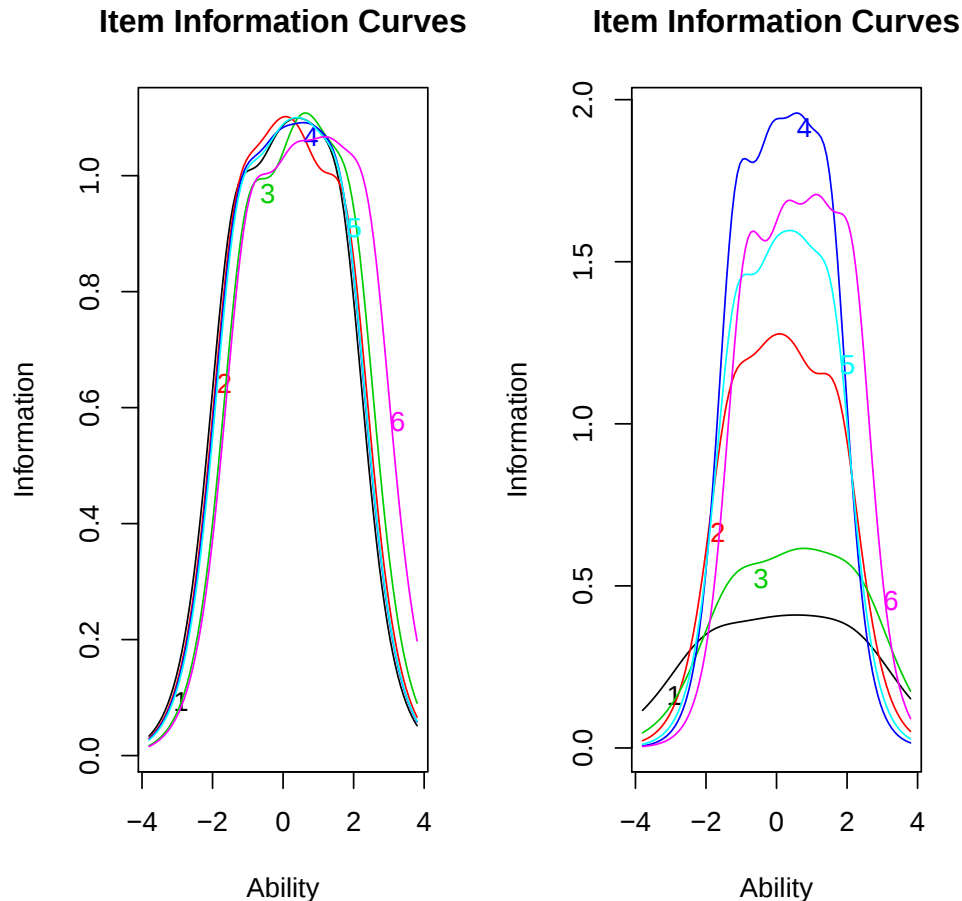


Figure 4.4: Item Information Curves, LOT-R, Sample Two, Split B for One Parameter GRM (left) and Two Parameter GRM (right)

Table 4.34 clearly shows that there were no obvious problems with this model. The parameter estimates are relatively low in comparison with other scales, suggesting that these items were easier to endorse.

The two parameter GRM (Appendix A, Table A.12, Page 206) suggested that the ability estimates have risen while the discrimination parameters have fallen for the majority of the items (except for 7).

The item information curves for Sample One suggest that Q4 (“I’m always optimistic about the future”) (Figure 4.4) could actually stand as a reasonably good proxy for

the entire scale.

Finally, we assess the performance of each of these models on unseen data.

Table 4.35: Performance of One and Two Parameter Graded Response Models on Unseen Data (Splits A and C)

	ErrorApproximation	Correlation
One Parameter	7.31	1.00
Two Parameter	12.99	0.99

As can be seen from Table 4.35, the one parameter model performed best on the unseen data, suggesting that the two factor model could not be used to reduce the scale.

An examination of the item ordering assumption showed that Q1 did not fit this model for Split C, and so was removed from the scale. This seems to suggest that the issues with Q1 are common across both samples, and that variability was the cause of the lack of this issue, in Split B. There were no failures of the monotonicity assumption, and thus the modelling could commence.

One and two parameter Graded Response Models were then fit to the items.

Table 4.36: Coefficient Estimates for LOT-R One Parameter Graded Response Model, Split C

	$\beta^1(se)$	$\beta^2(se)$	$\beta^3(se)$	$\beta^4(se)$	α
LOTRQ3	-1.158 (0.11)	-0.25 (0.1)	0.482 (0.11)	1.672 (0.11)	2.458 (0.09)
LOTRQ4	-0.812 (0.1)	0.257 (0.09)	0.738 (0.37)	1.733 (1.7)	2.458 (0.14)
LOTRQ7	-1.191 (0.15)	-0.156 (0.14)	0.542 (0.37)	1.605 (1.68)	2.458 (3.41)
LOTRQ9	-1.174 (2.39)	-0.041 (0.31)	0.624 (0.38)	1.736 (1.68)	2.458 (3.36)
LOTRQ10	-0.673 (0.13)	0.339 (0.13)	0.989 (0.13)	2.015 (0.13)	2.458 (0.13)

Table 4.36 shows the estimated difficulty parameters for the one parameter Graded Response Model. It can be seen that the discrimination parameter is quite high, and that the most difficult question is Q10 which is “Overall, I expect more good things to happen to me than bad”. The “easiest” question is Q7, which is one of the negatively phrased questions, suggesting that the two of these questions would be enough to garner a rough estimate of ability from participants.

When the estimates for the two parameter GRM (Appendix A, Table A.13, Page 207) were examined, it can be seen that Q7 is the most discriminating question, while still having the lowest ability estimates, suggesting that it is a very good question for separating out optimism and pessimism. Q10 is still the most difficult, but not as discriminating as Q7.

Finally, the performance of these two models is tested against unseen data.

As can be seen from Table 4.37, the one parameter model provides the best fit to the unseen data.

Table 4.37: Performance of LOT-R Split C One and Two Parameter Graded Response Models on Unseen Data

	ErrorApproximation	Correlation
One Parameter	2.52	1.00
Two Parameter	7.91	1.00

4.9.4 Back testing - LOTR

Each of the splits showed that a one factor solution provided the best fit to the data, but the IRT approach from Split B showed that Q1 did not provide a good fit to the scale. Therefore, for the back testing process the two models examined were both one factor models, but the second one had Q1 removed from the scale.

Table 4.38: Performance of Full and Reduced Model for LOT-R on Sample One Data

base	comparison	ep	minus2LL	df	AIC	diffLL	diffdf	p
LOTRBackTestReduced		10	2305.50	5.00	313.24			
LOTRBackTestReduced	LOTRBacktestFull	12	2450.89	9.00	39.05	145.39	4.00	0.00

Table 4.38 shows the results of the full and reduced one factor model on the Sample One Data. It can clearly be seen that the reduced model provides a better fit, even though the RMSEA is quite high (0.09). All other fit indices are quite good, however. Therefore, this reduced scale will be used in the experimental portion of the research.

Next, the fit of the different IRT models was examined using the same approaches as before.

Table 4.39: Performance of LOT-R Split C One and Two Parameter Graded Response Models on Unseen Data

	ErrorApproximation	Correlation
One Parameter Split B	9.37	1.00
One Parameter Split C	7.41	1.00

As can be seen from Table 4.39, the model developed on Split C with only five items fit the data from Sample One better than did the full model. This further shows both the utility of this approach, and its usefulness in selection of models for the experimental portion of the research.

4.10 Health, Optimism and Mindfulness Analysis

Following the investigation of the scales from a psychometric point of view, the remaining hypotheses of the study were addressed.

Table 4.40: Spearman Correlations between Scale Totals across both Samples

	physfun	rolelim	rolelimem	energyfat	emwellbeing	socialfunctioning	pain	generalhealth	mindfulness	optimism
physfun	1.00	0.27	0.13	0.22	0.18	0.22	0.28	0.31	0.16	-0.14
rolelim	0.27	1.00	0.33	0.31	0.24	0.37	0.33	0.30	0.25	-0.15
rolelimem	0.13	0.33	1.00	0.45	0.56	0.55	0.20	0.24	0.37	-0.29
energyfat	0.22	0.31	0.45	1.00	0.65	0.50	0.31	0.42	0.44	-0.43
emwellbeing	0.18	0.24	0.56	0.65	1.00	0.61	0.25	0.34	0.49	-0.54
socialfunctioning	0.22	0.37	0.55	0.50	0.61	1.00	0.36	0.32	0.39	-0.33
pain	0.28	0.33	0.20	0.31	0.25	0.36	1.00	0.36	0.21	-0.14
generalhealth	0.31	0.30	0.24	0.42	0.34	0.32	0.36	1.00	0.29	-0.32
mindfulness	0.16	0.25	0.37	0.44	0.49	0.39	0.21	0.29	1.00	-0.36
optimism	-0.14	-0.15	-0.29	-0.43	-0.54	-0.33	-0.14	-0.32	-0.36	1.00

As can be seen from Table 4.40, the optimism hypothesis was not supported. Contrary to predictions, optimism was negatively correlated with health (using Spearman's correlation coefficient). In fact, optimism correlated negatively with all of the other totals, suggesting that something unusual happened in the sample. This finding is discussed further throughout this chapter, Chapter 7 and Chapter 8.

Participants of both genders showed the relationship in the same direction, with participants reporting greater health reporting less optimism. The result is a general trend across all subgroups divided by college, suggesting that it is the result of a general pattern across the sample rather than being driven by some small number of aberrant observations. Indeed, when participants were stratified by College of study, the same trend was apparent suggesting that the relationship was consistent across all sub-groups.

It can be seen from Table 4.40 that the relationship between health and mindfulness was positive, and of the same magnitude at that observed between health and optimism.

Gender did not appear to have a substantial effect on mindfulness totals, although it is interesting to note that the range of health scores reported was much greater in the female participants.

MAAS scores were associated with greater health as expected, as can be seen from Table 4.40

4.10.1 Sample One, Relationships between Health, Optimism and Mindfulness

Table 4.41: Coefficients for General Health Regression Model with predictors chosen by Stepwise Selection on unseen data, Sample One

	B	SE(B)	β	t	Sig.(p)
(Intercept)	25.13	11.26		2.23	0.03
optimism	-1.68	2.17	-0.09	-0.78	0.44
energyfat	0.27	0.11	0.28	2.42	0.02
pain	0.12	0.09	0.14	1.29	0.20
rolelim	0.12	0.06	0.21	1.88	0.06

Stepwise regression was used for the selection of predictors. Typically, such an approach is not optimal in research, as the p-values are biased by the search process. This is also a problem when variables are manually added or removed from a model, as in standard exploratory practice. In this research, stepwise regression methods were used appropriately, as all p-values are reported on an independent test set, which means that they are unbiased. Additionally, stepwise provides a mathematical justification (in this case, the AIC) for the addition or removal of variables from the

model, which was useful in this research given the surprising, non-theoretically predicted results. Another benefit of this approach (in contrast to other selection methods such as lasso and ridge regression) is that stepwise provides a p-value for each coefficient, in line with APA conventions. As can be seen from Table 4.41, optimism, role limitations, pain and energy/fatigue were retained as predictors in the model. None of the variables were significant, with the exception of energy/fatigue and role limitations.

4.10.2 Regression Model Predictions

The model developed by stepwise selection was applied to the new data.

Table 4.42: Coefficients for Stepwise Selected Model on Test Data, General Health (Sample 2A)

	B	SE(B)	β	t	Sig.(p)
(Intercept)	5.10	0.52		9.85	0.00
generalhealth	-0.01	0.01	-0.17	-1.48	0.15
energyfat	-0.01	0.01	-0.16	-1.10	0.28
emwellbeing	-0.02	0.01	-0.37	-2.50	0.02

As can be seen from Table 4.42, the stepwise model has an extremely significant coefficient for emotional well being, and non-significant coefficients for energy fatigue and general health. The coefficient signs are all in line with those from study one. This confirmation between the model fit on a subset of Sample One and that fit on a subset of Sample Two increases the confidence that we should have in this solution.

4.10.3 Regression Analyses, Sample Two

Given the correlation matrix reported shown above in Table 4.40, regression analyses were run on the three major variables (General Health, Mindfulness and Optimism) to determine which other variables were involved in the effect.

4.10.4 Optimism

Table 4.43: Coefficients for Stepwise Selected Regression Model for Optimism

	B	SE(B)	β	t	Sig.(p)
(Intercept)	4.89	0.29		17.13	0.00
generalhealth	-0.01	0.00	-0.12	-2.10	0.04
Age	-0.01	0.01	-0.11	-2.07	0.04
pain	0.00	0.00	0.09	1.60	0.11
mindfulness	-0.09	0.05	-0.09	-1.60	0.11
emwellbeing	-0.02	0.00	-0.46	-7.86	0.00

In the training sample, the stepwise selected models kept three predictors. The results of the selected predictors on the test sample are shown in Table 4.43. General health is retained, suggesting that while it has an impact, they are moderated by the effect of emotional well being ($p \leq 0.01$). For further discussion of this relationship, see Chapters 7 and 8, as well as Section 6.7.

As has been shown above, the model showed that emotional well being and general health were important, and general health was not removed from the model, though it did not have a particularly large coefficient. Note that because the p-values reported for the regressions were calculated based on a subset of the data not used for selection of the predictors, these can be regarded as unbiased by the search process.

4.10.5 Mindfulness Regressions

A similar procedure as described above for optimism was employed in the mindfulness regressions. The results are shown below.

Table 4.44: Coefficients for Stepwise Selected Regression on Mindfulness

	B	SE(B)	β	t	Sig.(p)
(Intercept)	1.94	0.56		3.44	0.00
rolelim	0.01	0.01	0.16	1.27	0.21
rolelimem	0.00	0.00	0.06	0.48	0.64
emwellbeing	0.00	0.01	0.03	0.20	0.84
energyfat	0.02	0.01	0.39	2.89	0.01

As can be seen from Table 4.44, the stepwise selection process retained role limitation, emotional role limitations, emotional well being and energy fatigue, even though only emotional well-being remained significant. This model would seem to suggest that mindfulness is related to health through emotional well-being. This is interesting, in that in terms of face validity this would not be expected, although it is reported to be a relatively common outcome of MBSR.

4.11 Modeling the relationship between health, optimism and mindfulness

The next step in the analytic procedure was to examine the relationships between health, optimism and mindfulness somewhat more deeply, drawing on the work carried out so far with the data.

The first step was to assess the fit of the factor models and the regression models together, in a Structural Equation Model. This model was fit using the summary

scores. The regression model was formed by using the predictors from the stepwise regressions (on unseen data).

4.11.1 Optimism and Health

The most unexpected finding of this piece of research was the direction of the relationship between the LOT-R and the RAND MOS General Health scale. This section examines a number of different models, to determine which of the solutions is more likely, given the data.

The models used in this section were as follows:

- A model where both health and mindfulness, along with the other regression predictors, directly affect optimism;
- A model where the effect of health on optimism is mediated through the other predictor variables;

The regression model suggested that Age, Energy/Fatigue, Emotional Well Being, General Health and Mindfulness were associated with optimism, so these variables are included in the model.

Table 4.45: Comparison of Different Models between Optimism and Other Variables, Sample One

base	comparison	ep	minus2LL	df	AIC	diffLL	diffdf	p
LOTRRandDirect		15	2216362.61	2139.00	2212084.61			
LOTRRandDirect	LOTRRandIndirectMAAS	11	2216394.94	2143.00	2212108.94	32.33	4.00	0.00
LOTRRandDirect	RandIndirectEmWellBeing	12	1637243.78	2142.00	1632959.78	-579118.83	3.00	1.00
LOTRRandDirect	LOTRIndirectEnergyFatigue	11	2139147.38	2143.00	2134861.38	-77215.22	4.00	1.00
LOTRRandDirect	LOTRIndirectGenHealth	11	2106433.19	2143.00	2102147.19	-109929.42	4.00	1.00
LOTRRandDirect	LOTRHealthMediated	15	2216361.86	2139.00	2212083.86	-0.75	0.00	1.00

The different models for how optimism is related to the other variables (directly versus mediated) were compared and the results are shown in Table 4.45. As can be seen from this table, the best fitting model was that where Emotional Well Being mediated the relationship between optimism and health. Note that all of these models fitted quite poorly, so this finding needs to be tested on the second sample.

4.12 Sample Two Structural Equation Models

Table 4.46: Comparison of Models for Health, Sample Two

	chisq	df	pvalue	cfi	rmsea	rmsea.ci.lower	rmsea.ci.upper
Direct	0.00	0.00		1.00	0.00	0.00	0.00
EmWBMediated	299.20	5.00	0.00	0.81	0.21	0.19	0.23
MindfulnessMediated	375.84	5.00	0.00	0.60	0.24	0.22	0.26

As can be seen from Table 4.46, the best fitting model is one in which Emotional Well Being mediates the relationship between health and optimism, in line with the results shown in Sample One.

4.13 Discussion

A number of interesting findings, some expected and others not, have arisen from this study. Firstly, in contrast to published research, optimism and health were negatively associated. This relationship appears to be mediated through Emotional Well Being, which may point to some earlier work suggesting that the optimism-health connection was mediated through negative affect. This result would seem to contribute some new information to this debate.

Secondly, the factor structures for the RAND-MOS, the MAAS and the LOT-R were more or less confirmed, and more importantly from the perspective of this thesis, reduced versions of these scales were developed which were more predictive on new data.

4.13.1 Optimism and Health

The most striking finding is the large negative correlation between optimism and self reported health. The sample in this study is quite large, so the result is unlikely to be a statistical fluke. That being said, the result is problematic to explain, given the large amount of evidence of beneficial effects of optimism on health (Rasmussen, Scheier, & Greenhouse, 2009).

However, there have been some findings where higher optimism has not been associated with health outcomes. There are two major explanations for the curious and unexpected phenomenon, given the links established by a recent meta-analysis (Rasmussen et al., 2009). The first centres on the dimensionality of the optimism construct. Many believe that these results are caused by optimism self report measures being reflections of two interlinked constructs, optimism and pessimism (Herzberg, Glaesmer, & Hoyer, 2006) as established in a factor analytic study in a sample of over 46,000 participants. Some authors claim that the apparently contradictory results suggest that the pessimism part of the construct is the driver of the effects on health, and that the correlations between the two constructs decline with age. This viewpoint was partially supported by the recent meta-analysis which found that pessimism had a larger effect on health, though the difference between optimism and pessimism was not significant (Rasmussen et al., 2009). The other viewpoint argues that the effects of optimism on health are mediated by negative and positive affect, and that high levels of negative affect can either negate or reverse the

optimism-health link (Baker, 2007). The aforementioned Baker study found that the optimism health link was entirely mediated by negative affect. Nonetheless, the balance of the evidence suggests that optimism has beneficial consequences for health and healthy behaviours.

It is worth noting however, that the original literature of the beneficial effects of optimism on health focused on cellular immunity, which is obviously quite different from self reported health. A problem with this explanation is that reports in the literature indicate that the optimism-health link is larger when self report methods are used (Rasmussen et al., 2009).

A more likely explanation (especially given the results reported in Chapter 7) is that this unexpected finding was caused by the order in which the measures were administered. In the experimental study, these measures were administered in opposite order, and the correlations became positive. This would suggest that something about answering the questions of the MAAS impacted the scores on the LOT-R. Given that the MAAS is framed in terms of mindlessness, then it seems plausible that answering these questions and becoming aware of how they do not often meet these standards depressed scores on the optimism measure. This theory is additionally supported by the results of the SEM model which suggested that Emotional Well-Being (which could be regarded as a proxy for optimism, given the similarity of the questions) mediated this relationship.

4.13.2 Mindfulness

Another interesting finding which arose from this research is the impact of mindfulness scores on other health variables. MAAS scores correlated positively with all of the health sub-scales, very significantly in the case of emotional well-being. This may suggest that brief mindfulness interventions may be of use for improving overall population health, both physical and mental. That being said, it is important to note that the issues surrounding the mindfulness construct itself and its relations with mindfulness meditation practice need to be resolved before such strong conclusions can be drawn.

Another fascinating finding in this research is the strong negative correlation between mindfulness and optimism. This study appears to have been the first to assess these constructs using self-report measures, and this finding was not expected to occur. MBSR programs have been found to increase optimism in a number of studies (Carson et al., 2004), but our results seem to show that mindfulness and optimism may be inversely related - but see discussion above for a counterpoint.

There are a number of reasons why this could be so. Optimism is defined as generalised positive outcome expectancies about the future, while mindfulness is

defined as non-judgemental awareness of the content of thoughts. It seems plausible that increased mindfulness could lead persons to become less optimistic, as their new-found awareness of their own thought patterns and behaviours makes them aware that events have not always worked out well. This increased awareness could temper future assessments of the future, and decrease optimism as measured by the Life Orientation Test Revised. However, work in the experimental sample casts new light on this theory, and is discussed further in Chapter 7.

4.13.3 Psychometric Analysis

This study also confirms the proposed one factor structure for the MAAS, in line with previous research. This sample also appears to show that the LOT-R can be modelled without loss of information with just one factor. A replicable and parsimonious 8 factor structure for the RAND MOS was demonstrated, and our results cast further doubt on the notion that these factors are uncorrelated.

It is worth noting that in all cases, parallel analysis did not provide a good measure of the best number of factors to retain. For all three measures, the MAP criterion provided a more accurate metric. This may have resulted as parallel analysis procedures tend to sample from a normal distribution, and this condition was not met for any of our variables. The use of multiple decision criteria on a regular basis in factor analytic research would help to understand which method suits a particular application best.

Additionally, this study found that a combination of IRT analyses and factor analytic analyses provided the best fitting models for both the MAAS and the LOT-R. In both cases, a model fitted using the scales developed from the use of Mokken analysis and IRT fit indices provided better performance on unseen data than did the models developed exclusively through factor analysis.

This provides evidence that the combination of these two methods is more useful than either of them alone, and it has allowed for the building of a useful model tailored to the population under study in this thesis which can be used in the experimental portion of the research. The success of these models is discussed further in Chapter 7.

In conclusion, this portion of the thesis aimed to develop tailored measures for important covariates of the relationship between placebo and implicit and explicit expectancies, and this has been achieved. Additionally, this portion of the research sheds a new perspective on the relationships between self-reported health, optimism and mindfulness.

Chapter 5

Treatment Credibility Questionnaire

5.1 Introduction

Expectancies are considered to be at the core of the placebo response (Montgomery & Kirsch, 1997). However, currently they are assessed using very simple methods which fail to capture the multidimensional nature and changes over time in these important constructs (Stone et al., 2005). The most typical method of measuring treatment-related expectancies involves a one item measure “How much better do you expect to feel after this treatment”, which is answered on an eleven point scale (Morton et al., 2009; Martin-Pichora, Mankovsky-Arnold, & Katz, 2011). These expectancy measures normally have a correlation with placebo response of approximately 0.3, which equates to only 10% of the variance in response to placebo (Whalley et al., 2008).

5.1.1 Current Expectancy Measurement

The issues surrounding current expectancy measurement have been discussed in Chapter 2, especially Section 2.3.2, and the purpose of this portion of the research was to develop and test a more detailed measure of treatment expectancies which could then be used in the experimental portion of the research.

A review of expectancy studies showed that of 16 studies, only two shared a common expectancy question (Myers et al., 2008). This suggests that there is a need for a standardised measure of patient expectancies in the study of placebo effects and expectancy effects more generally.

This chapter describes an adaptation of the Credibility/Expectancy Questionnaire (Deville & Borkovec, 2000) to measure treatment expectancies in a student sample, and reports on its psychometric properties and correlation with other measures across two samples taken from the University population of staff and students.

The CEQ was chosen as its design lends itself to adaptation to different forms of treatment, which means that if the adaptation proves valid and reliable, then it can easily be used in other studies. Additionally, the CEQ was found to have both cognitive and emotional factors by the authors in the paper introducing the technique (Deville & Borkovec, 2000). These two factors have been shown by a systematic review to be important factors in response to placebo (DiBlasi et al., 2001). This theoretical link strengthened the argument for the use of this measure.

The CEQ was altered for this study, as this cross-sectional study did not administer any treatments to the sample of participants. Instead, participants were asked to imagine that they had been suffering from pain, and that a doctor had recommended

a particular treatment and were then asked to rate the credibility of a number of different methods for reducing their pain (described further in Section 5.3.3).

In this sense, what was measured were the self-reported treatment expectancies of the sample(s). These expectancies are relevant both because they may relate to the placebo response to treatments, and additionally because they may shape the choice of treatment (Bausell et al., 2005; Benedetti, 2005) (c.f. Section 5.1.2).

Finally, as this thesis aimed to develop and test an implicit measure to assist in the prediction of placebo, the use of an adequate self-report scale was important in that it allows for the incremental validity of any implicit measure to be assessed. The CEQ was used as a base as this set of questions has been shown to have validity in related areas, and this meant that it was a less risky process to follow when testing a new measure.

Current or former use of the treatments measured was not a requirement for the completion of the measure, but this is typical for research in placebo, and indeed, clinical trials more generally (Kirsch, 1985, 1997). The CEQ has been used in a family therapy setting for children with conduct disorder (Nock, Ferriter, & Holmberg, 2007), in the assessment of relational therapy for Vietnam war veterans and their spouses (G. J. Devilly, 2002) and in the treatment of PTSD (G. J. Devilly & Huth, 2008). It has not been used in any non-clinical samples, and this is its first adaptation for the study of treatment expectancies.

5.1.2 Beliefs About Treatment

One facet of recent research into both placebo and complementary and alternative medicine (CAM) has been some awareness that perceived treatment assignment may be as influential in outcomes as observed treatment assignment. In a re-analysis of four large RCT's into acupuncture analgesia, Bausell found that while actual assignment to treatment (verum acupuncture) was not associated with improved outcomes, beliefs about having received the verum treatment were significantly associated with better outcomes (Linde et al., 2007).

Additionally, in another study investigating the impact of acupuncture analgesia on recovering from dental pain following surgery, a similar effect was observed (Bausell et al., 2005). This suggests that beliefs about treatment assignment are as important as is treatment assignment itself (Benedetti, 2005). The relevance to this work is that these studies have shown that belief in receiving a real treatment is associated with greater improvement. Additionally, studies on the placebo effect have demonstrated that belief in the veracity of a substance can have profound treatment effects.

Therefore, by measuring treatment related expectancies more precisely it may become possible to predict which (if any) of a set of treatments a patient will respond to best.

5.1.3 Links to theory

There are a number of theoretical perspectives around how expectancies relate to placebo effects. The most prevalent in the field is the response expectancy perspective of Kirsch (Kirsch, 1985, 1997) which argues that expectancies are directly responsible for observed placebo responses, and that all other variables are mediated by them. Kirsch's theory draws from some of Albert Bandura's work on self-efficacy (Bandura, 1977), but is quite distinct in that response expectancies are regarded as a separate type. Crow *et al* (Crow *et al.*, 1999) take a somewhat different approach to expectancies, breaking them down into specific kinds of process expectancy (those expectancies developed by patients themselves) and positive and negative outcome expectancies which are either created or enhanced by the health provider.

For the purposes of this thesis, and in line with the theory around placebo discussed in Chapter 3, especially Section 3.2.6, treatment expectancies are regarded as outcome expectancies formed around particular sets of treatments which are then either enhanced or denigrated by subsequent learning. Essentially, expectancies are valid for a particular point in time, and are more reflective of state, and are amenable to change through either manipulations (Kirsch & Weixel, 1988) or through experience (Stewart-Williams & Podd, 2004). Indeed, this perspective is entirely compatible with the position of Stewart-Williams that conditioning is a means of creating new expectancies. Crucially for this theory, expectancies are regarded as having physiological impact which shapes the experience of a given treatment, which then impacts further response to a particular treatment or treatments in general.

5.1.4 Pain Treatments

In general, treatments for pain are available in multiple different forms. Pills are commonly available as an over-the-counter supplement, creams are available for muscle pain, and injections of morphine or other painkillers are provided in a health-care setting. In one study, the most common reason for taking some form of medication was headache (9%, N=2590 in a US sample) suggesting that these forms of treatment are routinely used (Kaufman, Kelly, Rosenberg, Anderson, & Mitchell, 2002). Another recent randomised survey of California residents determined that the point prevalence of some form of chronic pain condition in the general population was 49% (47.0%-51.0%, N=3243), which would seem to suggest that the majority of the population suffers at least some pain once per month (Ohayon & Schatzberg, 2010). These findings would seem to suggest that pain-related treatment expectancies are important to an understanding of pain treatment and the experience of pain more generally. This is another benefit which a useful measure for these extremely common treatment-related expectancies would provide.

Complementary and alternative methodologies are sometimes used to treat pain. In some cross-sectional research, 23% of patients with chronic pain reported using CAM in the previous twelve months (McEachrane-Gross, Liebschutz, & Berlowitz, 2006). In this study, CAM use was associated with more education and higher income. In another study of primary care participants undergoing treatment with opioids, 44% of the sample (N=908) reported using at least one CAM treatment in the past twelve months. CAM use was associated with age, gender, pain severity and income. This study also found that over half of the sample reported some benefits from these treatments.

For these reasons, both conventional and alternative treatment stimuli were used in this study. Additionally, it was thought that there might be some useful information conveyed by the relative ordering of conventional and alternative treatment expectancies, both within and between participants.

5.1.5 Aims and Objectives

The overall aim of this study was to develop and test a measure of treatment expectancies in a pain context which could then be used in the experimental study. To this end, the following was done:

- The reliability & general psychometric structure (Study 1) of the adapted questionnaire was tested using factor analysis and IRT;
- The validity of the original factor structure was assessed (Study 1) using both factor-analytic (FA) and item response theory (IRT) methods
- The best factor structure from Study 1 was replicated on a sub-sample of the Study 2 data
- The remainder of the data from Study 2 was then examined to assess if an exploratory analysis provided different results from Study 1;
- The convergent validity of the new measure was assessed using the BMQ (Study 2);
- The impact of demographic (age, gender, self-reported health, income and experience with the treatments concerned) variables on the newly developed measure (Study 2) was assessed.

5.2 Methodology

5.3 Design

This portion of the research was designed in the following fashion. In the first study (N=299), the measure was tested for reliability and to ensure that the factor structure from the previous research replicated in this new setting, and with some changes made.

The second study (N=1329) was conducted to validate the factor structure from the first study, and to assess convergent validity through the use of the BMQ. This study included two more forms of alternative treatment (Homeopathy and Reiki) to assess if a common factor structure could also be found for the alternative treatments.

Additionally, the second sample collected demographic statistics around income, health and familiarity with all of the painkilling treatments used in the major instrument, to assess if these would have the impacts observed in previous research.

This design had the advantage of both providing evidence for reliability of the instrument across Studies One and Two, and also providing independent evidence for the validity of the predicted factor structure in Study Two. The second, larger study provided links to important potential moderators of the treatment expectancies and assessed convergent validity using the BMQ.

5.3.1 Sampling

The sampling for the Treatment Credibility Questionnaire was conducted in two rounds, as the instrument was developed using one sample, and then confirmed and revised with another. The first sample was sent to a random subset of students at the University in February 2010 (N=2500), to which 299 students responded (RR=12%). This data was analysed over the next two months, and following this, a revised version of the questionnaire was sent to a random sample of both staff and students at the University (N=10000), to which 1329 people responded (RR=13.3%).

5.3.2 Participants

The mean age of the participants in Study 1 was 22.61 (SD=6.5), the median age was 21 and the range of reported ages was 10-56, which suggests some data entry errors on the part of the participants. The sample was 68.8% female, and 83.4% consisted of undergraduate students.

The sample size collected for Study Two was 1329, however, only 566 completed all questions in the two scales. The analysis was carried out on those participants who

had provided data for all questions, with each random split being inclusive of the missing values. Of the participants who filled out all questions, 71.5% were female ($N=396$), 82.4% were undergraduates, and the mean age was 22.96 ($SD=6.3$).

5.3.3 Measures

The Treatment Credibility Questionnaire was adapted from Devilly and Borkovec's CEQ, as described above. This measure has six questions. The scale was developed to assess expectancies around psychological therapies. In the original form, the measure has four questions (Set I) and another set of two questions (Set II) (Devilly & Borkovec, 2000). Questions 4 and 6 were scored on an 11 point scale, while the other questions were scored on a nine point scale. A number of changes were made to this instrument for use in this research. Firstly, the scale was changed to a 1-5 scale, to simplify the scoring and standardise it. Secondly, the six questions in each condition were prefaced by a statement that read: You have been suffering pain for a number of days. You go to the doctor, and he/she suggests you try X; where X is one of Pills, Creams, Injections or Acupuncture in Study One, or Pills, Creams, Injections, Acupuncture, Homeopathy and Reiki in Study Two.

The questions were as follows:

1. How logical does the therapy offered to you seem?
2. How successful do you think this treatment will be in reducing your symptoms?
3. How confident would you be in recommending this treatment to a friend?
4. How much improvement in your symptoms do you think will occur?
5. How much do you really *feel* that therapy will help you to reduce your symptoms?
6. How much improvement in your symptoms do you really *feel* will occur?

Of these, the first three questions were found to cluster together (using principal components analysis) in a credibility factor, while the second three questions were found to make up an expectancy factor (Devilly & Borkovec, 2000). Note that this was despite their arbitrary division into Set I and II, and was replicated across all three samples (where $N \approx 100$ for each of the samples).

In light of evidence that the IAT correlates better with explicit measures when they are relative measures (i.e. a feeling thermometer) rather than absolute measures, a non-relative measure was adapted for this research, to maximise the incremental variance explained by each measure.

5.3.3.0.1 Beliefs About Medicine Questionnaire Additionally, the Beliefs about Medicine Questionnaire (BMQ) was administered to all participants in Sample Two. This is an instrument developed to assess the beliefs of chronic pain patients regarding their medicines (Horne, Weinman, & Hankins, 1999). The BMQ General is an eight item measure which has been found in previous research to have two factors, and which was developed using principal components analysis (Horne et al., 1999). Only the General sub-scale was used for this research, as the Specific sub-scale was not appropriate for the non-clinical sample used in this research, which is a decision that has also been taken by other researchers (Mårdby, Åkerlind, & Jørgensen, 2007).

This instrument was included in order to validate the TCQ by means of correlations with this similar measure. The BMQ has been extensively used in clinical research, and has been shown to predict adherence to medication in chronic pain, depression and with general pharmacy patients in Sweden (Horne et al., 1999; C. Brown et al., 2005; Mårdby et al., 2007). Specifically, Maardby *et al* found that the General-Harm sub-scale of the questionnaire showed a significant relationship to adherence in the overall pharmacy-using population. This would seem to indicate that the measure (at least the General part) is useful with non-clinical populations.

5.3.4 Analysis

Study One was treated as an overall training and development sample (most appropriate given that the measure was changed as a result of these methods), and the results tested on a subset of Study Two.

Data for Study Two was split into four parts of approximately 300 observations each, in line with the procedure described in Chapter 3.

Firstly, the data was checked for errors in entry or recording using summary functions and plots. Following this, the question responses were recoded according to the instructions for use. Following this, the summary scores were calculated. Next, summary statistics and characteristics of the data were reported. Following this, a rank-order correlation matrix for the data was calculated and analysed.

Simple reliability analyses were carried out on the scales themselves. Following this, parallel analysis, the MAP criterion and the scree plot were used to estimate the number of factors which could be extracted from the data. Further analyses were carried out in line with the methods described in Chapter 3, Sections 3.4.1.4 and 3.4.2, in line with the aims and objectives laid out in Section 5.1.5. In brief, the first study used all of the data available to fit psychometric models and tested these on a subset of Study Two data. Ten-fold cross-validation was used for all regression models, and all p-values reported are from the test set.

5.4 Study One - Results

As described above, the aim of this Study was to ensure that the altered measure possessed face validity and was reliable and that the sub-scales correlated with each other in the expected fashion.

5.4.1 Demographic Statistics

The first step was to look at the scores on the response variables regarding the different forms of treatment.

In Table 5.1 the summary statistics for the mean credibility scores for each of the four treatment options are shown. It can be seen that credibility of analgesic creams was lowest, even lower than that of acupuncture. Note that the minimum score of zero was from participants who only completed some of the questions.

Table 5.1: Summary Statistics, Total Credibility Scores

	Mean	SD	Median	Min	Max
PillTot	3.48	0.96	3.83	0.00	5.00
CreamTot	2.87	1.01	3.00	0.00	5.00
InjTot	3.65	1.04	3.83	0.00	5.00
AcuTot	3.12	1.19	3.25	0.00	5.00

Next, the differences in credibility totals between Genders was examined.

There were no significant differences in median Pill scores that can be attributed to the gender of the respondents. The variability was larger within the male responses, a pattern which was repeated for Cream and Injection credibility totals also.

However, the acupuncture credibility scores showed a different pattern. The mean credibility totals were very different for females and males. A Kruskal Wallis examined whether this visual difference was backed up with a formal analysis. The results tended towards significance ($p=0.09$). The pattern of larger variability in the responses of men was apparent.

There were differences in the credibility scores of various treatments across the colleges of study. These differences were only significant for the Cream ($p = 0.0158$) and Injection ($p = 0.0168$) credibility scores.

5.5 Psychometric Analyses

5.5.1 Reliability Analysis

The reliability of this scale was assessed using Cronbachs α which is the most commonly used measure of reliability in psychological research.

This analysis was performed on the sample, and the mean alpha was equal to 0.9, which is well above the threshold used for survey research (0.7) and would, on the basis of this sample, qualify as a clinical instrument, for which the threshold is 0.9. This reliability analysis did not suggest that any of the items should be removed from the scale.

Table 5.2: Item reliability Statistics for the TCQ 1. Overall alpha of the scale was 0.92, following reverse scoring of the Acupuncture Items. Reverse scored items are indicated by the minus symbol. Raw/standard alpha (Cronbach's alpha), G6: Guttman's Lambda 6 (squared multiple correlation) S/N: Signal/Noise ratio, alpha_se: Standard error of alpha estimation

	raw_alpha	std.alpha	G6(smc)	average_r	S/N	alpha se
Pill1	0.89	0.89	0.97	0.27	8.49	0.01
Pill2	0.89	0.90	0.97	0.27	8.58	0.01
Pill3	0.89	0.90	0.97	0.27	8.62	0.01
Pill4	0.89	0.89	0.97	0.27	8.48	0.01
Pill5	0.89	0.90	0.97	0.27	8.58	0.01
Pill6	0.89	0.90	0.97	0.27	8.53	0.01
Cream1	0.89	0.90	0.97	0.28	8.78	0.01
Cream2	0.89	0.90	0.97	0.27	8.72	0.01
Cream3	0.89	0.90	0.97	0.28	8.76	0.01
Cream4	0.89	0.90	0.97	0.27	8.69	0.01
Cream5	0.89	0.90	0.97	0.27	8.66	0.01
Cream6	0.89	0.90	0.97	0.27	8.62	0.01
Inj1	0.89	0.89	0.97	0.27	8.49	0.01
Inj2	0.89	0.89	0.97	0.27	8.45	0.01
Inj3	0.89	0.90	0.97	0.27	8.64	0.01
Inj4	0.89	0.89	0.97	0.27	8.39	0.01
Inj5	0.89	0.89	0.97	0.27	8.37	0.01
Inj6	0.89	0.89	0.97	0.27	8.38	0.01
Acu1	0.91	0.91	0.98	0.32	10.67	0.01
Acu2-	0.90	0.90	0.97	0.29	9.43	0.01
Acu3-	0.90	0.90	0.97	0.29	9.38	0.01
Acu4-	0.90	0.90	0.97	0.29	9.42	0.01
Acu5-	0.90	0.90	0.97	0.29	9.40	0.01
Acu6-	0.90	0.90	0.97	0.29	9.35	0.01

Table 5.2 shows the α coefficient for each of the items, along with the Guutman Lambda 6 statistic, which measures the amount of variation in an item which can be accounted for by a linear regression of all the other items. Standardised alpha is based

on the correlations rather than the covariances, and as such as typically a much more useful measure.

5.5.2 Factor Analysis

There were two aims of this factor analysis on Study One. The first was to examine if the predicted factor structure from the development of the scale held. This was that there would be three expectancy and three credibility questions. This analysis was carried out on all four methods of treatment individually. The next step, following this analysis, was to examine the intra-treatment factor structure.

Interestingly enough, the proposed factor structure from earlier work replicated for the Pill credibility items, but not for the others. There was a two factor structure for the Injection credibility items, but the first factor carried all of the items except for item 3. For Cream credibility scores and Injection credibility scores only one factor was found from both a parallel analysis and from fitting a two factor model. All of these models had excellent fit indices, with high NNFI (approximately 0.95), and so can be regarded as fitting the data relatively well. Possible explanations for this finding are examined in the Discussion.

The next step in the analysis was an assessment of the number of factors suggested by both the parallel analysis and MAP criterion for the entire scale. The results of the parallel analysis and scree plot suggested four factors. The MAP criterion suggested that six factors should be extracted. Both of these factor solutions were extracted and examined based on fit indices and interpretability.

As can be seen from Table 5.3 the four factor structure is extremely interpretable, with each group of six questions loading highly on its own factor. This fits with the expectations prior to the research.

The Acupuncture factor does not correlate highly with the other three factors, but they do correlate reasonably well with one another. This makes clear that the distinction between the western and alternative treatment modalities was apparent to the participants.

This four factor solution explains 74% of the variance, which is extremely high for a psychological self report scale.

Next, the six factor solution was examined for interpretability, and the results are shown in Table 5.5. The loadings on Factors 5 and 6 (tentatively titled Credibility and Expectancy, respectively) are clustered around questions 1 and 3 (for credibility) and 2 and 4 (for expectancy). Note that the loadings are not particularly high for the Acupuncture questions, reinforcing the notion that these factors are measuring matters common to the conventional treatments. As can be seen in Table 5.5 the four

Table 5.3: Four Factor Solution, TCQ 1 Oblimin rotation

	Acupuncture	Cream	Inj	Pills	Communalites
Pill1	-0.03	0.1	0.02	0.7	0.58
Pill2	-0.04	0.06	0.03	0.68	0.54
Pill3	0.02	-0.02	0.05	0.73	0.57
Pill4	-0.01	0.02	0.08	0.75	0.66
Pill5	0.01	0.01	-0.01	0.8	0.63
Pill6	0.01	-0.03	-0.04	0.9	0.74
Cream1	0	0.82	-0.06	0.04	0.67
Cream2	-0.03	0.91	-0.02	-0.04	0.78
Cream3	0.06	0.81	0.02	0.02	0.7
Cream4	-0.01	0.91	0.05	-0.07	0.8
Cream5	0.03	0.8	0.01	0.08	0.72
Cream6	-0.01	0.83	0.02	0.06	0.75
Inj1	-0.05	0.03	0.78	-0.02	0.63
Inj2	-0.04	0.05	0.88	-0.09	0.73
Inj3	0.02	-0.07	0.78	0.02	0.59
Inj4	0.01	0.03	0.9	-0.03	0.8
Inj5	0.04	0	0.87	0.06	0.81
Inj6	0.01	-0.04	0.82	0.12	0.78
Acu1	0.87	0.05	0.01	0.01	0.77
Acu2	0.94	0.03	-0.03	0.01	0.91
Acu3	0.93	-0.01	0.01	-0.03	0.86
Acu4	0.96	0.01	0.03	-0.03	0.92
Acu5	0.97	-0.01	0	0	0.93
Acu6	0.96	-0.04	-0.01	0.02	0.92

Table 5.4: Four Factor TCQ correlations

	Acupuncture	Cream	Injections	Pills
Acupuncture	1.00	0.15	-0.15	-0.05
Cream	0.15	1.00	0.42	0.50
Injections	-0.15	0.42	1.00	0.67
Pills	-0.05	0.50	0.67	1.00

factors from the original solution still have the highest loadings, but there are a number of lower loadings on the fifth and sixth factors.

As can be seen in Table 5.7, the six factor model appears to fit the data much better, which is unexpected. Possible explanations are discussed below.

5.6 Item Response Theory Analyses

An IRT based item-selection procedure showed that the overall scale appears to divide into two scales, one for the conventional items and another for the acupuncture items. Therefore, these two sets of items will be analysed separately. Pill2 was the only item which did not meet the assumptions of the model in that it violated the

Table 5.5: Six Factor Solution, TCQ 1, Oblimin Rotation

	Acu	Cream	Inj	Pill	Cred	Exp	Communalites
Pill1	-0.03	0.04	-0.04	0.47	0.61	0.11	0.73
Pill2	-0.01	0.02	0.02	0.57	0.25	0.43	0.67
Pill3	0.02	-0.08	-0.01	0.51	0.61	0.08	0.7
Pill4	0.01	0.02	0.1	0.71	0.07	0.36	0.75
Pill5	-0.01	0.03	0.01	0.8	0.15	-0.13	0.76
Pill6	0.01	-0.01	-0.01	0.88	0.14	-0.02	0.83
Cream1	-0.01	0.78	-0.1	-0.1	0.34	0	0.76
Cream2	-0.02	0.89	-0.02	-0.07	0.02	0.23	0.83
Cream3	0.05	0.77	-0.02	-0.09	0.3	-0.03	0.76
Cream4	0	0.9	0.07	-0.01	-0.14	0.15	0.83
Cream5	0.01	0.83	0.04	0.14	-0.06	-0.2	0.83
Cream6	-0.03	0.88	0.05	0.14	-0.11	-0.18	0.88
Inj1	-0.06	-0.01	0.73	-0.14	0.34	0.04	0.73
Inj2	-0.01	0.04	0.87	-0.11	-0.02	0.42	0.85
Inj3	0.01	-0.1	0.73	-0.07	0.32	-0.05	0.68
Inj4	0.02	0.05	0.9	0.04	-0.15	0.31	0.86
Inj5	0.03	0.03	0.87	0.14	-0.07	-0.02	0.88
Inj6	0	-0.01	0.82	0.19	-0.04	-0.06	0.86
Acu1	0.87	0.04	0	-0.02	0.07	0.02	0.78
Acu2	0.95	0.02	-0.03	-0.01	0.02	0.04	0.91
Acu3	0.93	-0.02	0	-0.04	0.03	-0.01	0.86
Acu4	0.96	0.01	0.04	0	-0.05	-0.02	0.92
Acu5	0.97	-0.01	0.01	0.03	-0.04	-0.03	0.94
Acu6	0.96	-0.04	-0.01	0.03	-0.02	-0.01	0.92

Table 5.6: TCQ Six Factor Solution Correlations

	Acu	Cream	Inj	Pill	Cred	Exp
Acu	1.00	0.16	-0.15	-0.05	-0.01	-0.07
Cream	0.16	1.00	0.37	0.42	0.33	0.10
Inj	-0.15	0.37	1.00	0.57	0.37	0.01
Pill	-0.05	0.42	0.57	1.00	0.24	-0.05
Cred	-0.01	0.33	0.37	0.24	1.00	0.31
Exp	-0.07	0.10	0.01	-0.05	0.31	1.00

Table 5.7: CFA on TCQ Models, Sample One

base	comparison	ep	minus2LL	df	AIC	diffLL	diffdf	p
TCQ4		48	1374.04	252.00	1013.98			
TCQ4	TCQ6	58	913.84	242.00	573.77	-460.20	-10.00	

item ordering assumption (assumption of monotonicity). Therefore, this item was removed from the scale before further analyses.

The alternative TCQ (i.e. the Acupuncture items) showed no violations of the item ordering or monotonicity assumptions, and so no items were removed.

The next step was to fit a series of graded response models.

Table 5.8: Coefficients for the TCQ 1 Conventional Scale, One Parameter Graded Response Model

	$\beta^1(se)$	$\beta^2(se)$	$\beta^3(se)$	$\beta^4(se)$	α
Pill1	-2.113 (0.17)	-1.029 (1.23)	-0.639 (0.07)	1.301 (0.12)	1.884 (12.39)
Pill3	-1.578 (0.15)	-0.745 (1.23)	0.019 (0.12)	1.559 (0.15)	1.884 (0.07)
Pill4	-2.221 (0.15)	-1.33 (0.07)	-0.663 (0.25)	1.564 (15.11)	1.884 (0.2)
Pill5	-2.299 (0.39)	-1.039 (0.18)	-0.358 (0.25)	1.58 (0.07)	1.884 (0.17)
Pill6	-2.182 (0.07)	-1.06 (0.14)	-0.368 (0.27)	1.694 (0.16)	1.884 (0.48)
Cream1	-1.576 (0.14)	-0.25 (1.3)	0.334 (0.07)	2.082 (0.15)	1.884 (0.47)
Cream2	-1.756 (0.12)	-0.425 (1.28)	0.342 (0.14)	2.622 (0.14)	1.884 (0.07)
Cream3	-1.381 (0.12)	-0.155 (0.07)	0.778 (7.65)	2.649 (0.15)	1.884 (0.17)
Cream4	-1.681 (0.13)	-0.464 (0.14)	0.445 (7.59)	3.015 (0.07)	1.884 (0.16)
Cream5	-1.661 (0.07)	-0.245 (0.12)	0.484 (7.59)	2.65 (0.21)	1.884 (0.15)
Cream6	-1.773 (0.18)	-0.445 (0.12)	0.44 (0.07)	2.728 (0.73)	1.884 (0.2)
Inj1	-1.917 (0.49)	-1.071 (1.14)	-0.657 (0.14)	0.963 (0.73)	1.884 (0.07)
Inj2	-2.402 (0.49)	-1.75 (0.07)	-0.913 (0.14)	0.778 (0.73)	1.884 (0.18)
Inj3	-1.66 (0.49)	-0.826 (0.15)	0.11 (8.44)	1.316 (0.07)	1.884 (0.17)
Inj4	-2.371 (0.07)	-1.489 (0.13)	-0.829 (8.4)	0.829 (0.14)	1.884 (0.16)
Inj5	-2.078 (0.19)	-1.365 (0.13)	-0.65 (0.07)	0.878 (12.44)	1.884 (0.16)
Inj6	-2.214 (1.24)	-1.424 (0.18)	-0.491 (0.15)	0.985 (12.39)	1.884 (0.07)

Table 5.8 shows the fit of the one parameter GRM. Note that the Cream credibility items are the most difficult, which fits with the lower mean credibility score on this variable observed in Table 5.1. The next step taken is to examine the fit of a more flexible model (a two parameter GRM, Appendix A, Table A.15, on Page 208), and assess whether or not this represents a useful improvement over the model above.

The result of this model comparison exercise showed that the 1 parameter model was a significantly better fit than the 2 parameter model ($p \leq 0.001$). The item information curves are perhaps more illuminating (Figure 5.1), as they show that the Injection items convey a large proportion of the information across the range when we allow separate discrimination parameters for individual items.

Next, the fit of the alternative TCQ (i.e. the acupuncture items) was examined using a series of graded response models.

Table 5.9: Coefficients for TCQ 1 Alternative Scale, One parameter Graded Response Model

	$\beta^1(se)$	$\beta^2(se)$	$\beta^3(se)$	$\beta^4(se)$	α
Acu1	-0.93	-0.29	0.49	1.94	3.65
Acu2	-1.12	-0.41	0.68	2.20	3.65
Acu3	-0.80	-0.18	0.77	2.04	3.65
Acu4	-0.99	-0.34	0.54	2.12	3.65
Acu5	-0.90	-0.31	0.63	2.09	3.65
Acu6	-0.96	-0.35	0.66	2.11	3.65

The model described in Table 5.9 has the following features. The discrimination parameter is reasonably high (reflecting that these items were endorsed less by many of the participants), and the parameter estimates are closely distributed around zero.

The model shown in Table 5.10 has a much lower estimate of the discrimination of the

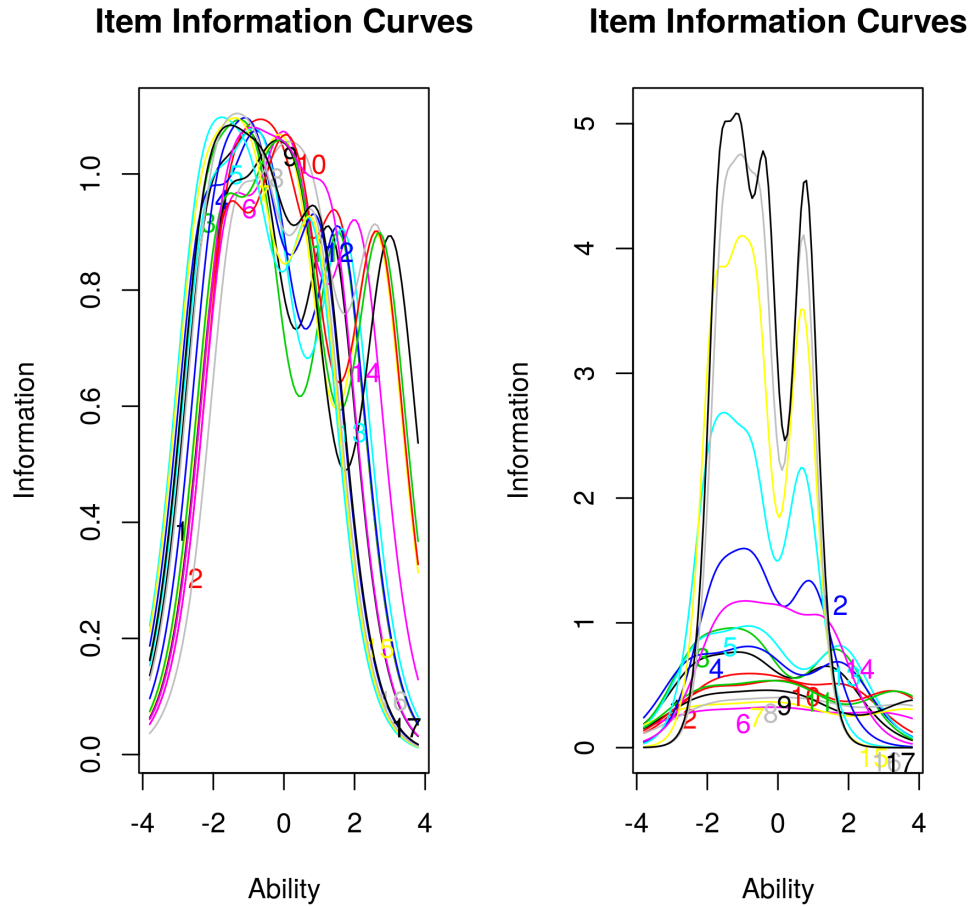


Figure 5.1: Item Information Curves, TCQ Conventional Scale, One Parameter GRM (top) and Two Parameter GRM (bottom)

Table 5.10: Coefficient Estimates for TCQ 1 Alternative Scale, Two Parameter Graded Response Model

	$\beta^1(se)$	$\beta^2(se)$	$\beta^3(se)$	$\beta^4(se)$	α
Acu1	2.05	1.28	0.41	-1.68	-3.05
Acu2	2.18	1.11	0.22	-1.86	-3.67
Acu3	1.85	0.97	-0.22	-1.61	-3.10
Acu4	2.08	1.03	0.19	-1.79	-3.81
Acu5	2.06	1.10	-0.07	-1.89	-3.22
Acu6	1.99	1.12	0.06	-1.71	-3.40

individual items, counterbalanced by higher ability estimates for all of the item difficulty parameters. Note that the one parameter GRM did not converge properly, and so standard errors are not reported.

5.7 Confirmatory Analyses

As in previous analyses (see Chapter 4), the second sample was much larger than the first, and therefore was split into a number of equal parts. In this case, the second

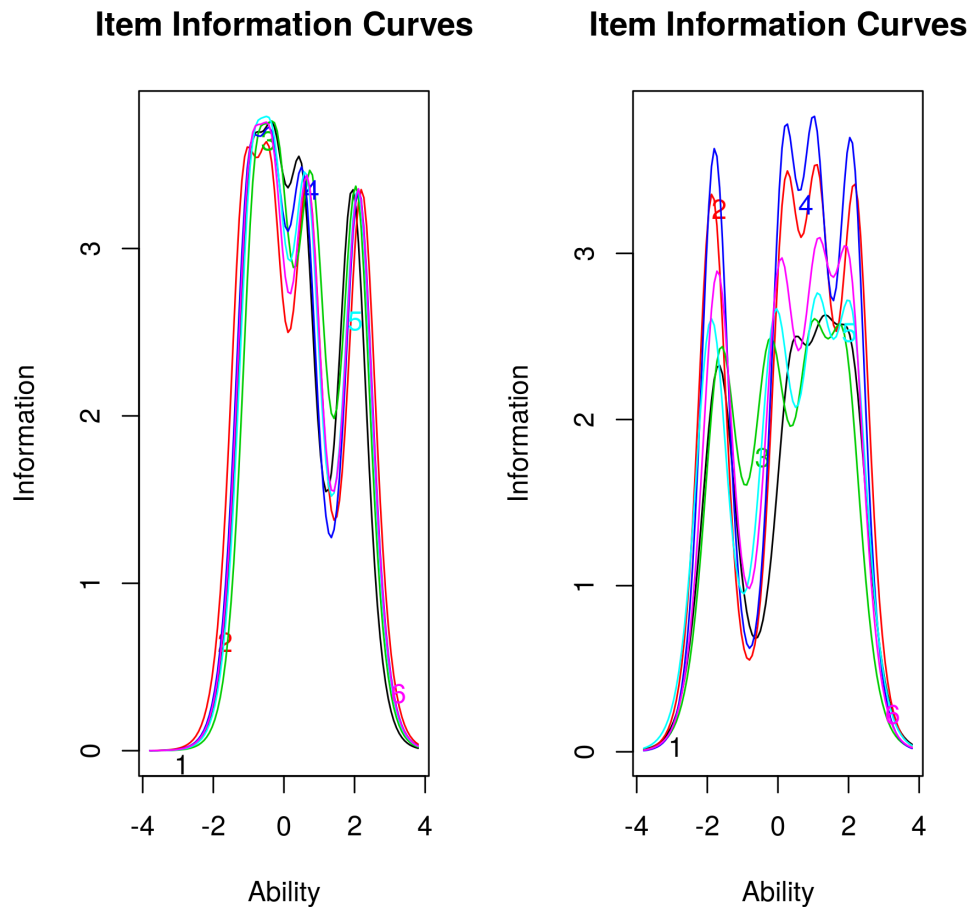


Figure 5.2: Item Information Curves, TCQ Alternative Scale, One Parameter GRM (top) and Two Parameter GRM (bottom)

sample (consisting of student and staff responses) was split into four equal sub-samples(A-D). The first of these (hereafter denoted as A), was then used as a test sample for the models developed on the data from Sample 1.

Table 5.11: Comparison of Four and Six Factor Models for the TCQ1 on Sample Two

base	comparison	ep	minus2LL	df	AIC	diffLL	diffdf	p
TCQ42		48	-265.64	252.00	1376.57			
TCQ42	TCQ62	54	-336.38	246.00	1317.83	-70.74	-6.00	

As can be seen from Table 5.11 the six model appears to fit the data best, even on this unseen data-set. This is somewhat unexpected, especially given the compelling reasons to believe in a four factor model. Further explanations are given in the Conclusion for Study One (Section 5.8).

The first step in examining the predictive power of the IRT models is to score the current subsample of the Study Two data (Split A) using the prior models. Then, the same model can be fit on this data alone, and the error of estimation between the two

processes can be calculated. Essentially, the root mean square error of approximation (RMSEA) was used to assess the usefulness of each of the models.

The Conventional TCQ items were assessed first.

Table 5.12: Comparison of One and Two Parameter Graded Response Models, TCQ Conv, Split 2A

	ErrorApproximation	Correlation
One Parameter GRM	17.60	0.98
Two Parameter GRM	40.39	0.86

As shown in Table 5.12, the one parameter model performed better on the unseen data than did the two parameter model (unlike the analysis using likelihoods which picked the more complicated model). The same pattern was apparent when the alternative scale was examined.

5.8 Conclusions & Next Steps

A number of issues have become clear from this first study validating the Treatment Credibility Questionnaire.

Firstly, the measure appeared to be quite reliable, and there were some interesting findings with respect to the measure. All of the treatments showed average levels of credibility, ranked from injections, pills, acupuncture and cream credibility scores. There was a pattern of greater variability in the responses of men to the questionnaire. This could be the result of less experience or from a wider range of beliefs in the population. If it is due to less experience, then it would be expected to see less variability in the responses of those individuals with more experience of these treatments. This theory was tested in Study Two.

With respect to the factor structure, the results of Horne (Horne et al., 1999) have not been confirmed in this adaptation, except for one of the measures. Pill credibility score showed the same two factor structure as did the CEQ in earlier work, but none of the other three treatment forms did. Indeed, the other three treatments showed a relatively clear one factor structure, though the results of a parallel analysis suggested that a two factor structure might be more appropriate for the Injection credibility scores.

It is important to consider what may be the reasons for the difference. One potential explanation is that this factor structure was obtained because of greater experience with painkilling treatments in pill form. If this is the case, then greater levels of experience with a painkilling treatment should correlate with the presence of this factor structure. Another possibility is that this simply due to chance, in which case this finding should not replicate in the next study.

With regards to the overall psychometric structure, the measure appears to show a strong pattern of one factor per treatment modality, but a six factor model fitted the training and test data better, suggesting that there are some higher-order correlations between some of the factors, which could probably be better accounted for using a hierarchical factor model.

Secondly, an item response theory analysis has suggested that while the Pills, Cream and Injection questions can be conceptualised with one latent trait, this does not appear to work for acupuncture. This would seem to suggest that ratings of complementary and western medicine are somewhat independent. The second version of this survey added two more sections on complementary therapies, Homeopathy and Reiki, in order to both balance the kinds of treatments, and to examine whether the assumption of bi-dimensionality would hold in another student and staff sample.

Very few demographic variables were collected in this survey. This limited what could be done with the ability estimates, and their relation to other factors. Therefore, for the second sample, more demographics were added. A question on experience with each of the treatments involved would seem like a useful check on what percentage of the respondents have experiential knowledge of these methods would prove useful. Given the cost of CAM and its unavailability to those of lower incomes, a question on income was also added in line with the results of prior research noted in the Introduction.

In addition, the General sub-scale of the Beliefs about Medicines Questionnaire was appended to V2 of the survey, in order to assess the extent to which the two instruments correlate (or fail to). This is a preliminary attempt to assess construct validity, which of course is but a prelude towards testing the instrument with clinical samples and in experimental settings.

5.9 Study Two - Results

Given the results of Study 1, this study added two more Complementary and Alternative Medicine (CAM) methodologies to the questionnaire, to determine if a higher order factor structure could be found which represented conventional and complementary methodologies. The hypotheses were as follows:

- The TCQ V2.0 will have six factors, one for each treatment modality;
- The BMQ will correlate negatively with the Pills, Cream and Injection totals, and positively with the Acupuncture, Homeopathy and Reiki totals;
- The TCQ will not fit an item response theory model well, and will need to be divided into TCQ Conventional (Pills, Creams, Injections) and a TCQ

Alternative (Acupuncture, Homeopathy and Reiki) in order to meet the assumptions of the model(s);

- Experience with a particular treatment will be correlated with higher ratings of its credibility;
- Experience with a particular treatment will be correlated with the presence of a two factor structure for that particular treatment.

5.9.1 Results

5.9.1.1 Descriptive Statistics

Table 5.13: TCQ2 Summary Statistics

	Mean	SD	Median	Min	Max
Age	22.96	6.34	21.00	17	65
Year	2.22	1.16	2.00	1	5
Income	1.08	0.38	1.00	1	4
Health	3.94	0.84	4.00	1	5
ExpPill	3.04	1.00	3.00	1	5
ExpCream	1.97	1.01	2.00	1	5
ExpInj	1.59	0.92	1.00	1	5
ExpAcu	1.26	0.67	1.00	1	5
ExpHom	1.29	0.73	1.00	1	5
ExpRei	1.18	0.62	1.00	1	5
Pill1	4.04	1.05	4.00	1	5

Many respondents stopped answering questions throughout the demographic variables. The reasons for this are unknown, but are probably not related to the design in any substantial way. The median age is 22 (SD=11.4), which shows that most of the sample was made up of students.

The Health variable is quite skewed, as most respondents rated their health as very good as good, as can be seen from the mean and the median (Table 5.13). The mean level of experience with a particular treatment was highest for pills and steadily declined for all of the other treatments.

The sample was 69.17% female, with 45 respondents not specifying their gender.

The next step in the analysis was visualisation of data using scatter-plot matrices and conditioning plots. Figure 5.3, the correlations between the scale totals can be seen. Notable are the high correlations between each of the conventional and alternative treatment methodologies, and their moderate correlations (in opposite directions) with the BMQ. Given the sample size, all correlations were significant at the $p \leq 0.001$ level.

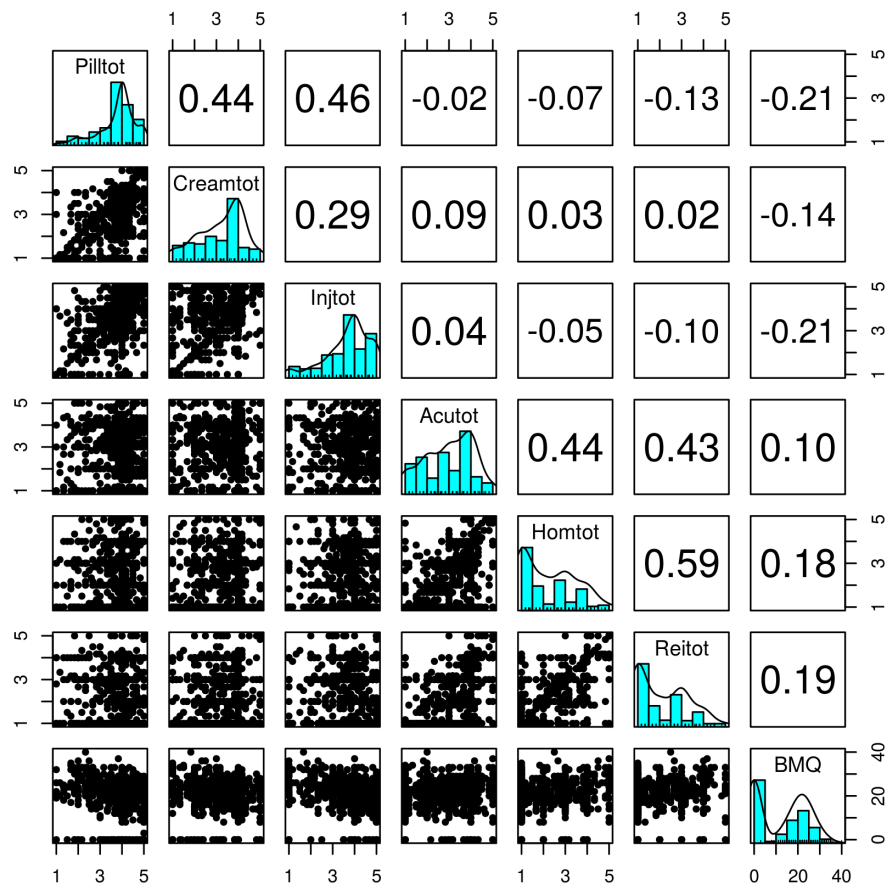


Figure 5.3: Scatter-plot Matrix, TCQ 2. The upper triangle represents correlations between the variables, the lower triangle shows scatterplots between them, and the diagonal shows a histogram of the values. From top to bottom: Pilltot - Pill Credibility, Creamtot - Cream Credibility, Injt看 - Injection Credibility, Acutot - Acupuncture Credibility, Homt看 - Homeopathy credibility, Reit看 - Reiki credibility, BMQ - BMQ Scores

The histograms for the alternative treatments are quite interesting, given that they appear to have a number of peaks with very little in between. This may reflect the polarised nature of the attitudes towards these treatments.

The sample scores showed some significant differences attributable to Gender. The cream, acupuncture, homeopathy and Reiki totals were significantly different in the sample, as was the Beliefs About Medicine scale total.

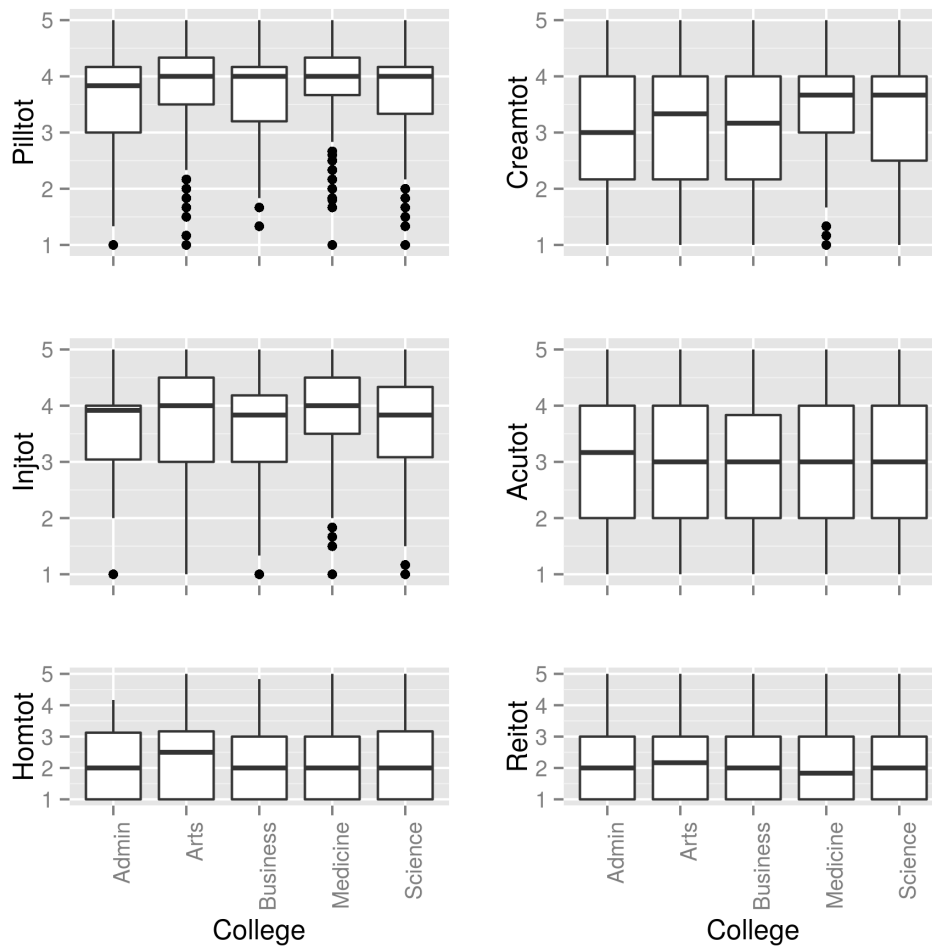


Figure 5.4: Credibility Scores by College

As can be seen from Figure 5.4, there were significant differences between the credibility totals in each college for the three conventional treatment modalities. This appears to be due to higher credibility totals amongst those respondents whose primary affiliation was with the College of Medicine and Health.

Next the range of experience scores with each of the treatment was examined.

Table 5.14: Experience with painkilling treatments by Gender

Gender	Pill	Cream	Inj	Acu	Hom	Rei
Not Reported	4.00	2.67	2.33	1.33	1.33	1.00
Female	3.03	1.83	1.52	1.28	1.32	1.19
Male	2.72	1.99	1.65	1.28	1.16	1.10

As can be seen from Table 5.14 females reported a higher level of pill experience, along with lower levels of cream and injection experience, and slightly higher levels of Homeopathy and Reiki experience.

Next, the response to the Income and Health questions were assessed.

Table 5.15: Income Levels for Study two

Income	Freq
Less than 20K	574
20K to 40K	105
40K to 60K	66
60 to 80K	25
80K+	11

As can be seen from Table 5.15, most participants reported earning less than 20,000 euros per year. This is unsurprising, as the majority of respondents were students.

Finally, the levels of General Health reported were investigated.

Table 5.16: Self Reported Health levels, Study Two

Health	Freq
Very Poor	2
Somewhat Poor	26
Average	185
Somewhat Good	353
Very Good	227

As can be seen from Table 5.16, the majority of participants tended to report their health as either Somewhat Good or Very Good, which is not particularly surprising, given the non-clinical nature of the sample.

Following these investigations of the descriptive qualities of the data, the next stage of the analysis is to examine the inter-correlations and latent structure behind the data through the use of factor analysis and IRT.

5.10 Treatment Credibility Questionnaire, Version 2

Splits B,C and D were used for this analysis, given the use of Split A as a testing set for the Sample One data (in Section 5.7).

Parallel analysis and the scree plot suggest that five factors seemed most optimal for this instrument. However, the MAP criterion disagrees, suggesting that seven factors might be more appropriate. To determine which of these criteria is correct, the five, six and seven factor solutions were examined for interpretability and then subjected to CFA on the other splits of the data. This procedure was repeated for each of the other splits, and the coefficients of each of the models on each split were averaged and are reported below.

5.10.1 Five Factor Solutions

The five factor solutions explained 79% of the variance in Split B and 81% in Split C and all of the communalities were extremely high. It was not suggested by any of the factor prediction methods in Split D.

The average of the two five factor solutions (Appendix A, Table A.14, Page 207) merged the Pill credibility and Injection credibility items together, and this pattern was repeated across both splits. Note that both the Cream and Homeopathy factors have slightly lower loadings than do the others, and the Pill & Injections factor is weighted more towards Injections.

Pillinj: This factor consists of the Pill and Injection items, and can therefore best be termed Pills and Injections.

Homeopathy: “Hom1”, “Hom2”, “Hom3”, “Hom4”, “Hom5”, “Hom6”. This factor maps exactly to the Homeopathy items, and is therefore given that name.

Reiki: “Rei1”, “Rei2”, “Rei3”, “Rei4”, “Rei5”, “Rei6”. Again, this factor maps exactly to the Reiki items and is named after them.

Cream: “Cream1”, “Cream2”, “Cream3”, “Cream4”, “Cream5”, “Cream6”. PA4 maps to the Cream items and so retains that name.

Acupuncture: “Acu1”, “Acu2”, “Acu3”, “Acu4”, “Acu5”, “Acu6”. PA3 maps to the acupuncture items and so is termed Acupuncture.

The Pills/Injections factor correlated highly with the Cream factor (as would be expected), while the Alternative factors correlate quite well with one another also. Of particular note is Acupuncture, which appears to occupy a middle ground between the other alternative methods and the conventional ones, as least as evinced by these correlation structures.

5.10.2 Six Factor Solutions

The next step was to fit six factor solutions over Splits B, C and D and comparing them to the five and seven factor solutions above and below respectively. Each factor solution was tested on unseen data.

As can be seen from Table 5.17, the six factor solution is relatively clean when averaged across all splits. The structure fits the predicted one extremely well. The six factors map exactly to the six treatment modalities. One small exception to this was that Pill1 loaded slightly onto the Reiki factor. However, in light of its strong loading on other factors, this was disregarded. The only two items for which the communalities are low are for Pill1 and Inj3, which may suggest some residual

Table 5.17: Average of Six Factor Solutions over Splits B, C and D

	Hom	Rei	Acu	Cream	Inj	Pill	Communalities
Pill1	0.1	-0.14	-0.09	0.04	0.11	0.62	0.57
Pill2	-0.04	-0.04	0.07	0.03	0	0.79	0.68
Pill3	0	-0.05	0	0.12	0.06	0.62	0.57
Pill4	-0.04	0.05	0.07	-0.01	0.06	0.79	0.7
Pill5	0.05	-0.02	-0.07	0.03	0.06	0.82	0.78
Pill6	0.01	0.04	-0.02	0.03	0.07	0.83	0.79
Cream1	0	-0.04	0	0.77	0.04	0.03	0.66
Cream2	-0.01	0.01	0.02	0.91	-0.01	0.01	0.84
Cream3	0.03	-0.05	0.02	0.81	0.03	0.01	0.71
Cream4	0.01	-0.01	0	0.96	-0.01	-0.03	0.89
Cream5	0	0.05	-0.02	0.91	0.01	0.04	0.88
Cream6	-0.01	0.02	0	0.94	0.01	0.02	0.91
Inj1	0	-0.02	-0.09	0.04	0.73	-0.06	0.52
Inj2	0	-0.03	0.06	0.03	0.82	0.07	0.8
Inj3	-0.02	0.03	-0.04	0.05	0.69	-0.02	0.49
Inj4	0.02	-0.03	0.04	-0.01	0.91	0.03	0.88
Inj5	0	0	0.02	0.02	0.9	0.06	0.9
Inj6	-0.01	0.03	0	0	0.9	0.06	0.87
Acu1	0.08	0.01	0.8	0.03	0	-0.06	0.75
Acu2	0	-0.01	0.94	0	-0.01	0.05	0.89
Acu3	0.03	0.04	0.82	0.03	0.04	-0.05	0.77
Acu4	-0.01	0.03	0.95	0	0.01	0	0.92
Acu5	0	0	0.96	-0.02	0.01	0.01	0.93
Acu6	0.02	0.03	0.94	0	0.01	0	0.95
Hom1	0.88	0.02	0.04	0.01	0.01	-0.03	0.86
Hom2	0.95	0.01	0.02	-0.02	-0.02	0.01	0.94
Hom3	0.87	0.05	0.03	0	-0.03	0.03	0.86
Hom4	0.98	-0.01	0.01	-0.02	0	0	0.96
Hom5	0.95	0.03	0	0.02	0	-0.01	0.95
Hom6	0.96	0.04	-0.01	0.01	0.01	-0.01	0.97
Rei1	0.09	0.86	0	0.03	-0.01	-0.03	0.88
Rei2	0.02	0.95	0	0	-0.01	0.01	0.95
Rei3	0.04	0.9	0.01	-0.01	0.02	0	0.89
Rei4	0	0.95	0.03	-0.01	-0.01	0.01	0.94
Rei5	0	0.97	0.02	0.01	-0.01	0	0.96
Rei6	0.02	0.96	0.01	-0.01	0.01	0	0.96

variance (see the seven factor solutions below). The three solutions explained from 79% to 84% of the variance.

Table 5.18 shows that the three conventional treatment factors and the three alternative factors correlated with each other but tended not to correlate highly with the factors of the opposite modality.

Table 5.18: Average of Factor Correlations, Six Factor Solution, TCQ2 (B,C,D)

	Rei	Hom	Acu	Cream	Inj	Pill
Rei	1.00	0.71	0.54	0.08	-0.10	-0.11
Hom	0.71	1.00	0.56	0.05	-0.07	-0.08
Acu	0.54	0.56	1.00	0.13	0.09	0.03
Cream	0.08	0.05	0.13	1.00	0.38	0.53
Inj	-0.10	-0.07	0.09	0.38	1.00	0.64
Pill	-0.11	-0.08	0.03	0.53	0.64	1.00

5.10.3 Seven Factor Solutions

The seven factor solution shown in Table 5.19 explained 86% of the variance in the sample, and the items were assigned to factors as described in the Table above. In essence, this solution was extremely similar to the six factor solution above, except that some of the Pill and Injection items loaded on the Credibility factor. The loadings tended to be highest for Q1 and Q3, and this factor would seem to account for the shared variance between these items not captured in the six factor solution (as noted above when discussing the low communalities for these items.)

The factor correlations are shown in Table 5.20, and follow a similar pattern to the six factor correlations. Note that the Pill and Injection factors correlate much more highly with the Credibility factor, which is a point that will be revisited below.

Figure 5.5 shows that the hierarchical solution breaks down into a conventional and alternative structure (both factors constrained to be equal). This suggests that modelling the data with two higher-order structures might be useful. However, this failed to provide a parameter-adjusted decrease in log-likelihood, and so was not pursued further.

5.10.4 Confirmatory Factor Analyses

Over all splits, the seven factor model provided a better fit to the out of sample data than did the 5 or 6 factor model. However, the difference in likelihood was not large, and given the strong theoretical reasons for preferring the six factor model, this was chosen.

5.10.5 IRT Analyses

The first step is to examine whether or not the TCQ should be split into separate scales, then check that the assumptions are met for each of these scales. Across all splits, the item selection procedure suggested that the scale should be split into conventional and alternative items. GRM models were fit across Splits B, C and D and the coefficients were averaged and reported here. The models were tested using a

Table 5.19: Average of Seven Factor Solutions over Splits B, C and D

	Pill	Cream	Inj	Acu	Hom	Rei	Cred	Communalities
Pill1	0.66	0.02	0.04	-0.07	0.09	-0.14	0.2	0.61
Pill2	0.77	0.05	0.02	0.08	-0.06	-0.03	-0.04	0.68
Pill3	0.71	0.08	-0.06	0.02	-0.02	-0.04	0.36	0.73
Pill4	0.75	0.01	0.11	0.07	-0.04	0.05	-0.07	0.7
Pill5	0.79	0.06	0.12	-0.07	0.05	-0.02	-0.13	0.8
Pill6	0.8	0.05	0.12	-0.02	0	0.05	-0.1	0.81
Cream1	0.04	0.75	0.02	0.01	0.01	-0.05	0.15	0.68
Cream2	0.01	0.9	0.03	0.02	-0.01	0.01	-0.06	0.85
Cream3	0.05	0.79	-0.06	0.04	0.02	-0.04	0.29	0.79
Cream4	-0.03	0.96	0.01	0	0.01	-0.02	-0.06	0.9
Cream5	0.03	0.91	0.05	-0.03	0	0.05	-0.08	0.9
Cream6	0.03	0.94	0.02	0	-0.02	0.03	-0.06	0.93
Inj1	-0.01	0.03	0.59	-0.06	-0.02	-0.01	0.37	0.61
Inj2	0.09	0.05	0.81	0.06	0.02	-0.06	-0.01	0.81
Inj3	0.04	0.01	0.55	-0.02	-0.05	0.05	0.54	0.73
Inj4	0.04	0.02	0.91	0.04	0.03	-0.06	-0.04	0.9
Inj5	0.07	0.05	0.89	0.02	0	-0.02	-0.03	0.92
Inj6	0.07	0.02	0.89	0.01	0	0.01	-0.04	0.89
Acu1	-0.05	0.03	-0.02	0.81	0.08	0.01	0.09	0.76
Acu2	0.04	0	0	0.94	0	-0.01	-0.02	0.89
Acu3	-0.04	0.01	-0.01	0.83	0.04	0.04	0.15	0.8
Acu4	0	0	0.02	0.95	-0.01	0.03	-0.05	0.92
Acu5	0.01	-0.02	0.03	0.96	0	0.01	-0.05	0.93
Acu6	0	0	0.01	0.94	0.02	0.03	-0.03	0.95
Hom1	-0.02	0.01	0	0.04	0.88	0.02	0.06	0.86
Hom2	0	-0.02	0	0.01	0.96	0.01	-0.01	0.94
Hom3	0.03	-0.01	-0.04	0.03	0.88	0.04	0.09	0.88
Hom4	0	-0.01	0.01	0.01	0.97	0	-0.03	0.96
Hom5	-0.01	0.02	0.01	0	0.95	0.03	-0.04	0.96
Hom6	-0.02	0.01	0.02	-0.01	0.96	0.04	-0.04	0.97
Rei1	-0.02	0.02	-0.03	0	0.09	0.86	0.05	0.88
Rei2	0.01	0	-0.01	0	0.02	0.96	0	0.95
Rei3	0	-0.01	0	0.02	0.05	0.9	0.07	0.9
Rei4	0.01	0	-0.01	0.03	0	0.95	-0.03	0.95
Rei5	0	0.02	-0.01	0.02	0	0.97	-0.02	0.96
Rei6	0	0	0.01	0.02	0.01	0.96	-0.04	0.96

factor-score based method described earlier on all of the data which had not been used to fit them.

5.10.6 Conventional Items

Next, the assumptions were checked for the conventional items, and there were no violations for either the IIO or monotonicity assumptions.

Table 5.20: Averaged Factor Correlations for the Seven Factor Solution, Sample Two

	Hom	Rei	Acu	Cream	Inj	Pill	Cred
Hom	1.00	0.71	0.53	0.10	-0.06	-0.08	-0.01
Rei	0.71	1.00	0.56	0.03	-0.10	-0.11	-0.06
Acu	0.53	0.56	1.00	0.13	0.09	0.03	-0.01
Cream	0.10	0.03	0.13	1.00	0.35	0.51	0.10
Inj	-0.06	-0.10	0.09	0.35	1.00	0.59	0.16
Pill	-0.08	-0.11	0.03	0.51	0.59	1.00	0.14
Cred	-0.01	-0.06	-0.01	0.10	0.16	0.14	1.00

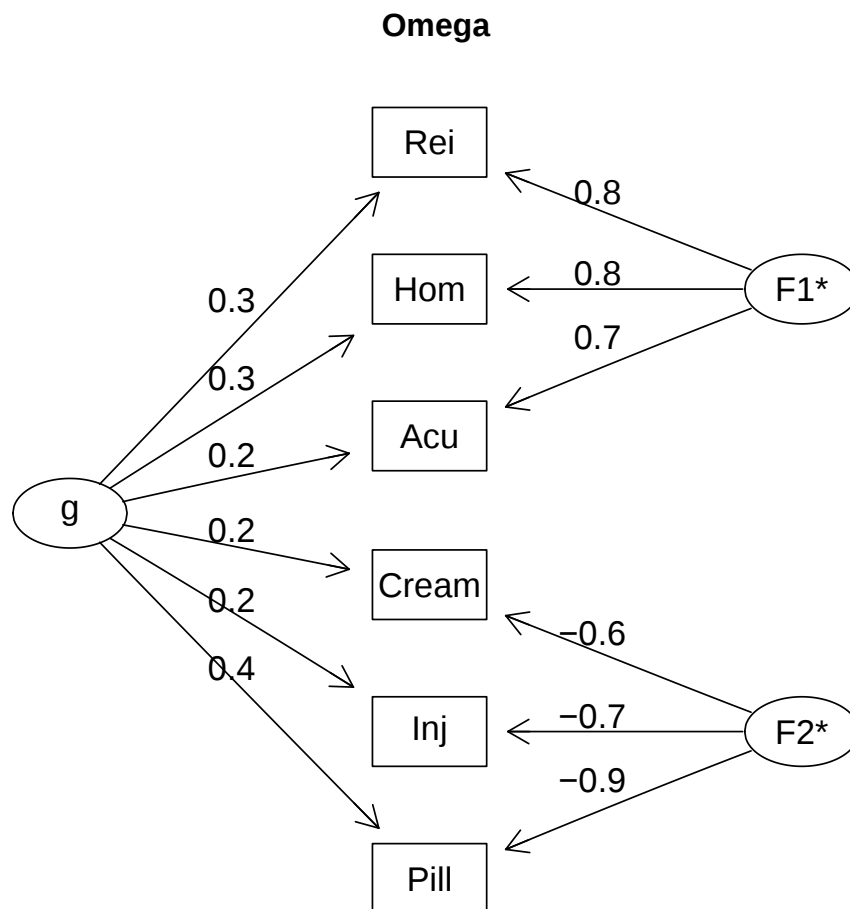


Figure 5.5: Hierarchical Factor Structure (Omega) for TCQ Data

Table 5.24 shows that the cream items were the least endorsed, especially the highest response category across all three splits. The averaged discrimination parameter is also quite high, at 2.12. This is probably due to a number of difficult items (the cream items) dragging up the estimated discrimination parameter.

Table 5.25 shows the estimated difficulty parameters for the two parameter GRM averaged across Splits B, C and D. Note that with this model, the cream credibility items have increased in difficulty, but the injection items have the highest

Table 5.21: TCQ Models from Split C tested against Splits A, B and D

base	comparison	ep	minus2LL	df	AIC	diffLL	diffdf	p
TCQ5notC		72	-1267.53	594.00	3242.91			
TCQ5notC	TCQ6notC	72	-2321.35	594.00	2290.18	-1053.82	0.00	1.00
TCQ5notC	TCQ7notC	75	-2547.22	591.00	2070.30	-1279.69	-3.00	

Table 5.22: TCQ 2 Models built on Split D, tested against Splits A, B and C

base	comparison	ep	minus2LL	df	AIC	diffLL	diffdf	p
TCQ6notD		72	-2253.35	594.00	2303.57			
TCQ6notD	TCQ7notD	75	-2442.69	591.00	2120.23	-189.34	-3.00	

Table 5.23: TCQ 2 Factor Solutions From Split B, Tested Against Splits A, C and D

base	comparison	ep	minus2LL	df	AIC	diffLL	diffdf	p
TCQ5notB		72	-747.47	594.00	3160.54			
TCQ5notB	TCQ6notB	72	-1785.01	594.00	2236.50	-1037.54	0.00	1.00
TCQ5notB	TCQ7notB	76	-1967.34	590.00	2062.17	-1219.87	-4.00	

Table 5.24: Item Difficulty Thresholds for TCQ Conventional items averaged over Splits B, C and D

	$\beta^1(se)$	$\beta^2(se)$	$\beta^3(se)$	$\beta^4(se)$	α
Pill1	-2.5 (0.24)	-1.42 (0.23)	-1.18 (0.23)	0.35 (0.22)	2.12 (0.09)
Pill2	-2.81 (0.28)	-1.71 (0.28)	-1.35 (0.28)	0.95 (0.28)	2.12 (0.09)
Pill3	-1.91 (0.16)	-1.07 (0.56)	-0.45 (0.56)	1.08 (0.56)	2.12 (0.09)
Pill4	-2.87 (0.29)	-1.55 (0.75)	-1.12 (1.1)	1.06 (1.1)	2.12 (0.09)
Pill5	-2.33 (0.2)	-1.24 (0.2)	-0.86 (0.97)	1.02 (0.96)	2.12 (0.09)
Pill6	-2.49 (0.23)	-1.29 (0.22)	-0.89 (0.22)	1.16 (0.36)	2.12 (0.09)
Cream1	-1.89 (0.17)	-0.89 (0.17)	-0.47 (0.16)	1.3 (0.17)	2.12 (0.09)
Cream2	-1.8 (0.16)	-0.77 (0.46)	-0.32 (0.46)	1.8 (0.46)	2.12 (0.09)
Cream3	-1.57 (0.14)	-0.59 (0.9)	0.17 (0.93)	1.82 (0.93)	2.12 (0.09)
Cream4	-1.79 (0.16)	-0.78 (0.14)	-0.22 (3.5)	1.87 (3.46)	2.12 (0.09)
Cream5	-1.75 (0.15)	-0.73 (0.14)	-0.17 (0.16)	1.85 (2.28)	2.12 (0.09)
Cream6	-1.69 (0.15)	-0.71 (0.15)	-0.14 (0.14)	1.94 (0.17)	2.12 (0.09)
Inj1	-1.66 (0.16)	-0.92 (0.6)	-0.6 (0.6)	0.52 (0.6)	2.12 (0.09)
Inj2	-2.2 (0.2)	-1.57 (7.18)	-1.03 (7.09)	0.4 (7.09)	2.12 (0.09)
Inj3	-1.52 (0.14)	-0.64 (0.12)	-0.06 (5.17)	0.96 (5.12)	2.12 (0.09)
Inj4	-2.15 (0.2)	-1.49 (0.19)	-0.89 (0.18)	0.51 (0.19)	2.12 (0.09)
Inj5	-2.16 (0.19)	-1.35 (0.2)	-0.78 (0.19)	0.59 (0.19)	2.12 (0.09)
Inj6	-2.03 (0.18)	-1.34 (0.79)	-0.77 (0.79)	0.65 (0.79)	2.12 (0.09)

discrimination, suggesting that these items are the best for distinguishing between high and low “ability” respondents. Standard errors are not shown as the models did not converge for Splits B and D.

Figure 5.6 shows that for the second sample, both the Cream and Injection items contributed a far greater proportion of the information when their slopes were allowed to vary independently. Interestingly enough, this pattern did not appear in Split B, suggesting that sampling variance was at play here.

Table 5.25: Item Difficulty Thresholds for Two Parameter GRM, TCQ Conv, Averaged over Splits B, C and D

	$\beta^1(se)$	$\beta^2(se)$	$\beta^3(se)$	$\beta^4(se)$	α
Pill1	-2.98	-1.54	-1.21	0.76	1.41
Pill2	-2.57	-1.59	-1.17	1.24	2.09
Pill3	-2.01	-0.95	-0.20	1.70	1.57
Pill4	-2.64	-1.34	-0.92	1.40	2.46
Pill5	-2.20	-1.01	-0.60	1.33	2.51
Pill6	-2.18	-0.99	-0.58	1.44	2.66
Cream1	-1.99	-0.71	-0.20	2.00	1.64
Cream2	-1.64	-0.53	-0.03	2.28	2.34
Cream3	-1.51	-0.34	0.58	2.59	1.79
Cream4	-1.58	-0.47	0.17	2.42	2.20
Cream5	-1.47	-0.42	0.16	2.26	2.42
Cream6	-1.40	-0.41	0.18	2.33	2.47
Inj1	-2.41	-1.07	-0.53	1.35	1.12
Inj2	-2.67	-1.78	-1.04	0.80	2.23
Inj3	-1.95	-0.48	0.48	2.06	1.21
Inj4	-2.89	-1.77	-0.84	1.21	2.42
Inj5	-2.40	-1.36	-0.57	1.14	2.68
Inj6	-2.35	-1.45	-0.64	1.20	2.52

5.10.7 Alternative Items

There were no violations of the invariant item ordering or monotonicity assumptions in this split for the alternative items. Two graded response models were fit to all three Alternative splits of the data, and the averaged results over all three splits are reported here.

Table 5.26 shows the averaged coefficients for the one parameter alternative item GRM. It can be seen that the Acupuncture items had the lowest difficulty thresholds, but all items had extremely high discrimination parameters.

The difficulty parameters for the averaged two parameter models increased across all items while the difficulty parameters have reduced significantly for most of the items. The Reiki items appear to be the most discriminating, while the Acupuncture items, while difficult appear to change less as the participant ability thresholds increase.

5.10.8 Testing the Models

Table 5.27 shows the average error across the three splits for one and two parameter GRM's for the conventional items. It can be seen that the one parameter model is clearly superior to the two parameter model, a pattern which was found across each individual split also.

Table 5.28 shows the results of a comparable process for the alternative items. Again,

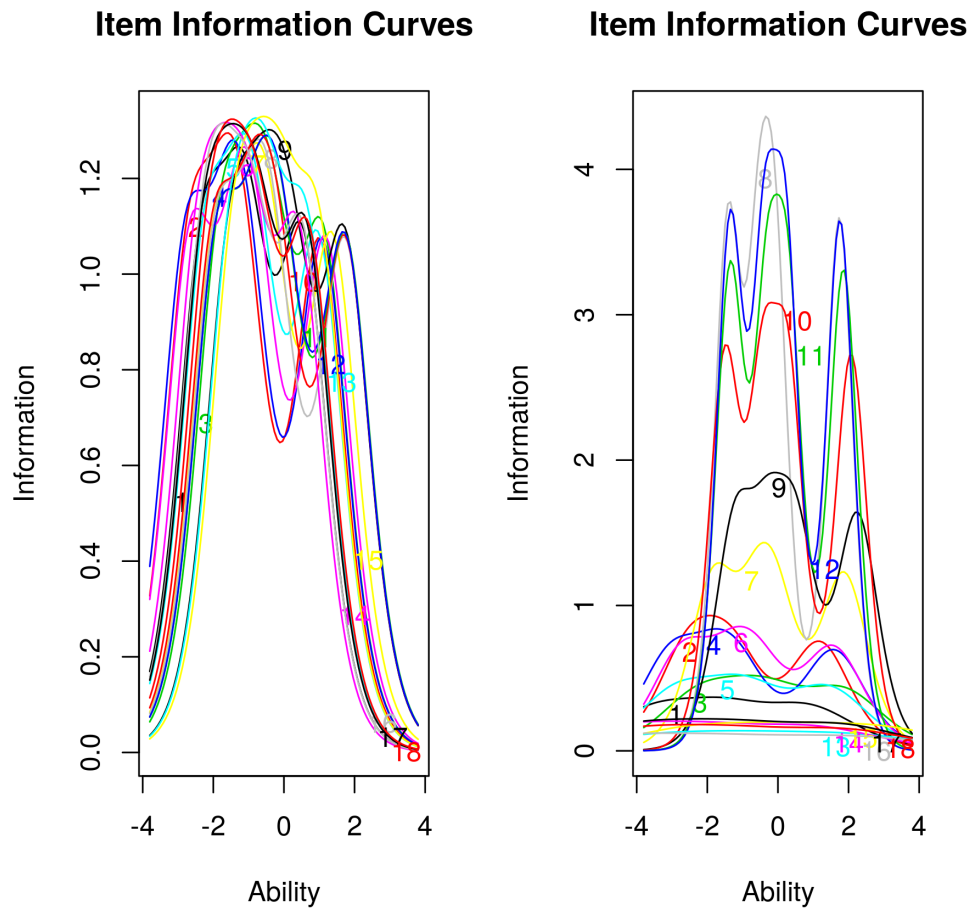


Figure 5.6: Item Information Curves, TCQ Conventional, Split 2, One Parameter GRM (top) and Two Parameter GRM (bottom)

the one parameter model proves superior, a pattern which was only repeated over Splits B and D.

5.10.9 Final Tests

The next stage in the analysis is to take the best of these models from each split and test them on the entire dataset. While this is, in a sense, analysing the data twice, it is the most practicable way in which to determine the best model for the entire sample¹.

The best fitting models on the out of sample data were as follows:

- Split B: TCQ7
- Split C: TCQ7
- Split D: TCQ7

¹due to the changes in the scale in Sample Two, a back testing strategy was not possible in this case

Table 5.26: Item Difficulty Thresholds for One Parameter GRM on the Alternative Items, averaged over Splits B, C and D

	$\beta^1(se)$	$\beta^2(se)$	$\beta^3(se)$	$\beta^4(se)$	α
Acu1	-0.44	0.00	0.16	0.58	15.82
Acu2	-0.37	-0.05	0.17	0.72	15.82
Acu3	-0.34	-0.00	0.24	0.58	15.82
Acu4	-0.34	-0.05	0.18	0.72	15.82
Acu5	-0.38	-0.04	0.20	0.68	15.82
Acu6	-0.34	-0.03	0.20	0.71	15.82
Hom1	-0.03	0.13	0.32	0.72	15.82
Hom2	-0.08	0.13	0.40	0.80	15.82
Hom3	-0.03	0.16	0.44	0.77	15.82
Hom4	-0.07	0.12	0.38	0.80	15.82
Hom5	-0.07	0.13	0.39	0.82	15.82
Hom6	-0.07	0.12	0.40	0.79	15.82
Rei1	0.04	0.17	0.43	0.79	15.82
Rei2	0.02	0.20	0.48	0.92	15.82
Rei3	0.05	0.19	0.51	0.84	15.82
Rei4	0.00	0.20	0.47	0.93	15.82
Rei5	0.01	0.18	0.48	0.92	15.82
Rei6	-0.01	0.18	0.50	0.93	15.82

Table 5.27: Average Error of Approximation for One and Two Parameter GRM's on the Conventional TCQ Across Splits B, C and D

	AverageError
One Parameter GRM	4.69
Two Parameter GRM	65.12

Table 5.28: Average Error of Approximation for One and Two Parameter GRM's on the TCQ Alternative items Across Splits B, C and D

	AverageError
One Parameter GRM	57.77
Two Parameter GRM	67.91

Each of these models will be run on the full dataset, and the results assessed. The model which fits best on this sample will be used to predict factor scores for the experimental data.

Table 5.29: TCQ2 Models from each split tested against the full dataset

base	comparison	ep	minus2LL	df	AIC	diffLL	diffdf	p
TCQ7notAAll		74	-2691.13	592.00	2822.45			
TCQ7notAAll	TCQ7notBAll	76	-2869.62	590.00	2647.96	-178.49	-2.00	
TCQ7notAAll	TCQ7notCAll	75	-2631.64	591.00	2883.94	59.49	-1.00	
TCQ7notAAll	TCQ7notDAll	75	-2732.56	591.00	2783.02	-41.43	-1.00	

As can be seen from Table 5.29, models B, C and D were identical and they appear to provide the best fit on all of the data.

5.11 Reducing the size of the TCQ

The TCQ Version 2 discussed in this chapter consisted of 36 questions. It was decided to reduce the number of items in the questionnaire in order to reduce the burden on participants, and to reduce the time needed to complete the experiment to less than one hour (in almost every case). Therefore, it was decided to reduce the TCQ to 18 questions, three for each methodology. Additionally, given that the measures across different forms of treatments needed to be comparable, this meant removing the same three questions across all treatment options.

This procedure was carried out using the IRT and Factor Analytic models built up in this chapter.

One interesting part of the factor analytical results reported above was that a seven factor solution was consistently a better performer on unseen data using SEM. This factor tended to consist of the pill and injection items 1 and 3. This would seem to suggest that these two items loaded on both the relevant treatment factor(s), and another overall factor that tended to be termed confidence in (conventional) treatment. Additionally, these two items tended to have lower factor loadings on their respective treatment factors. Therefore, it was decided to remove items 1 and 3 for each of the six sections of the TCQ before administering it to the experimental sample.

Figure 5.7 shows the estimated parameters for items one and three for conventional treatments. It can be seen that Item one consistently has extremely poor coverage over the range of abilities, given that only response thresholds 1, 4 and 5 cover any appreciable amount of the ability distribution. Perhaps surprisingly, items one and three for the alternative treatments did not have this problem, but overall they showed an extremely bi-modal split, with respondents either endorsing them whole-heartedly or rejecting them outright.

With the decision taken to remove the potentially problematic items one and three, the choice was between items two, four, five and six. Of these, items 2, 4 and 6 are coherent and cover overall expectancies, cognitive expectancies and emotional expectancies. Previous research has outlined these as two important factors of care (DiBlasi et al., 2001), and the three questions span a relatively broad span of ability ranges, and cover approximately 50% of the available test information between -3 and +3 (assuming that the ability scores are normally distributed).

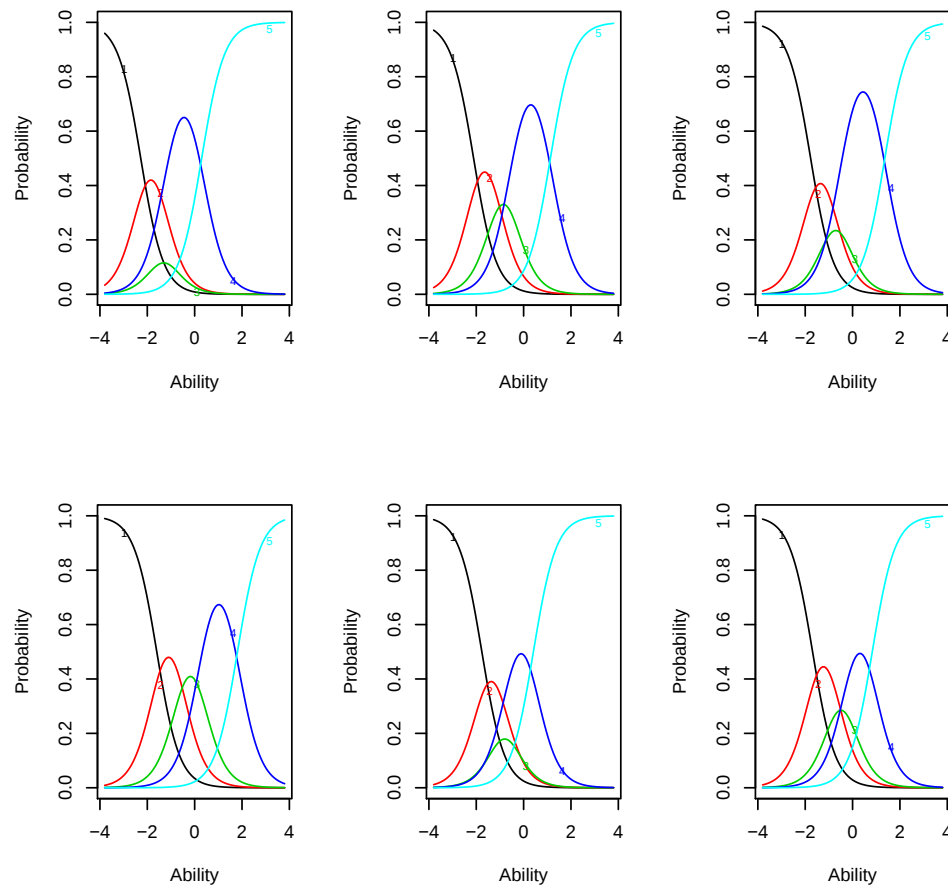


Figure 5.7: Information Curves for Items One and Three for TCQ-Conventional, Split A

5.12 Relationships between Credibility Totals and Other Variables

The next part of the analysis to be carried out was regression analysis. There were a number of demographic variables collected in the second sample, to allow for examination of their effects on the outcome variables (the treatment credibility scores). Gender, education, college of study or work, health, income and experience with the six forms of treatment were collected.

The method used was the following. Firstly, the sample was divided into training (70%) and test (30%) splits. Ten-fold cross-validation was performed on all training sets, and the best model(s) were then tested on the test set, which allowed accurate p-values and standard errors to be calculated on the test data.

From the model on the held out data, it can be seen that the results (Table 5.30) were highly significant $F(10,129)=54.76$, $p \leq 0.001$. The model had an adjusted R^2 of 0.51,

Table 5.30: Coefficients for Pill credibility Regression on Held Out Data

	B	SE(B)	β	t	Sig.(p)
(Intercept)	2.09	0.41		5.16	0.00
Creamtot	0.32	0.07	0.36	4.70	0.00
Injtot	0.34	0.06	0.43	5.46	0.00
Acutot	-0.00	0.06	-0.00	-0.03	0.97
ExpPill	0.07	0.06	0.09	1.21	0.23
ExpCream	-0.19	0.07	-0.20	-2.84	0.01
Age	-0.02	0.01	-0.11	-1.52	0.13

which equates to approximately 25% of the variance in Pill Credibility explained. The Cream and Injection variables were highly significant, while the Acupuncture and Experience with Pills variables did not pass the traditional significance filter.

Experience with creams did pass this filter however, which may be a function of increasing experience with multiple painkilling problems increasing their credibility.

The next step was to repeat this process using Cream total as a dependent variable.

Table 5.31: Coefficients for Cream credibility Regression on Held Out Data

	B	SE(B)	β	t	Sig.(p)
(Intercept)	-0.52	0.51		-1.00	0.32
Pilltot	0.55	0.10	0.49	5.55	0.00
Injtot	0.20	0.08	0.21	2.33	0.02
Acutot	0.17	0.06	0.18	2.73	0.01
ExpPill	-0.07	0.08	-0.07	-0.90	0.37
ExpCream	0.22	0.08	0.20	2.66	0.01
Health	0.09	0.08	0.08	1.18	0.24

The results for the held out data are shown in Table 5.31. The model was again significant, with an $F(6, 144) = 20.61$, and a p-value of $p \leq 0.001$. The adjusted R^2 for the model was equal to 0.4497. Pill credibility, experience with cream treatments, Injection and Acupuncture credibility remained significant,

Next, the same modelling procedure was applied to the Injection credibility scores.

Table 5.32: Coefficients for Injection credibility Regression on Held Out Data

	B	SE(B)	β	t	Sig.(p)
(Intercept)	0.17	0.45		0.38	0.70
Pilltot	0.81	0.08	0.68	9.81	0.00
Acutot	0.05	0.08	0.05	0.61	0.54
Reitot	0.03	0.07	0.04	0.46	0.65
ExpPill	0.05	0.07	0.06	0.80	0.43

The model results shown in Table 5.32 show that Pill credibility scores were excellent predictors, while none of the other selected predictors remained significant on new data. The model was independently significant: $F(4, 116) = 29.93$, $p \leq 0.0001$, and the adjusted R^2 was equal to 0.49.

Next, the same procedure was applied to the Acupuncture credibility scores.

Table 5.33: Coefficients for Acupuncture credibility Regression on Held Out Data

	B	SE(B)	β	t	Sig.(p)
(Intercept)	1.18	0.39		2.99	0.00
Creamtot	-0.06	0.10	-0.05	-0.58	0.56
Injtot	0.17	0.09	0.17	1.87	0.06
Homtot	0.31	0.12	0.32	2.63	0.01
Reitot	0.35	0.12	0.35	3.03	0.00
ExpAcu	0.22	0.13	0.13	1.68	0.10
ExpHom	-0.28	0.14	-0.16	-2.02	0.05

As shown in Table 5.33, only Homeopathy and Reiki, remained significant on the unseen data.

Following this, the same procedure was applied to the Homeopathy credibility totals.

Table 5.34: Coefficients for Homeopathy credibility Regression on Held Out Data

	B	SE(B)	β	t	Sig.(p)
(Intercept)	0.07	0.24		0.31	0.76
Acutot	0.30	0.07	0.29	4.45	0.00
Reitot	0.61	0.07	0.58	9.25	0.00
ExpAcu	-0.41	0.12	-0.19	-3.29	0.00
ExpHom	0.36	0.09	0.23	4.05	0.00
ExpCream	-0.01	0.07	-0.01	-0.19	0.85

Table 5.34 shows the estimated coefficients from the stepwise selected model on unseen data. It can be seen that Acupuncture and Reiki credibility remain significant, along with the relevant experience variables for each form of alternative credibility. The model has an R^2 of 0.62, which is extremely high.

The selection and testing procedure was then applied to the Reiki totals.

Table 5.35: Coefficients for Reiki credibility Regression on Held Out Data

	B	SE(B)	β	t	Sig.(p)
(Intercept)	0.49	0.52		0.94	0.35
Pilltot	-0.11	0.09	-0.08	-1.17	0.24
Creamtot	-0.01	0.08	-0.01	-0.18	0.86
Acutot	0.18	0.07	0.16	2.38	0.02
Homtot	0.67	0.07	0.67	9.76	0.00
ExpPill	-0.04	0.07	-0.04	-0.62	0.54
ExpRei	0.18	0.12	0.09	1.49	0.14
Health	0.01	0.08	0.01	0.14	0.89

Table 5.35 shows the performance of the stepwise selected coefficients on unseen data. It can be seen that the only coefficients which maintained significance were those for Homeopathy and Acupuncture credibility, while the R^2 was quite low, at approximately 0.4.

Interestingly, it appears that the experience with a particular form of treatment was only usefully predictive for the alternative treatments. Further comments on this are made in the Discussion.

5.12.1 Discussion

This chapter has demonstrated a number of matters with regards to this new instrument. Firstly, it appears to have the predicted factor structure, especially the reduced structure, with the cross-loading items removed.

Additionally, this factor structure has been validated by the use of unseen data to test and compare all models. This allows for maximum use to be made of the limited data collected in the experimental research, and is an approach which is not often employed within the field.

Secondly, this factor structure has been replicated over two independent samples. This is important as one sample may have been biased in some random direction, but two samples are much less likely to have been.

Thirdly, all the major hypotheses (from Study 2) have been confirmed. The measure appears to be stable, possesses good reliability, and seems to correlate in the expected ways with the BMQ General sub-scale. These findings would seem to suggest that this measure should prove useful in the assessment of treatment related expectancies in the experimental sample.

Interestingly enough, the hypothesis that experience with a particular treatment would correlate with experience of a particular treatment was only borne out for the alternative treatments. This data cannot be used to establish causality in this case, as both an experience to credibility and a credibility to experience causal chain are equally plausible.

The measure developed here was primarily aimed to be used in the experimental portion of the research, however given its design and its focus on expectancies, it may prove useful to researchers who are studying expectancies related to pain, or who wish to control for pre-existing expectancies in clinical or non-clinical research.

Further development of this instrument could profitably examine the impact of a different person framing the health treatment. It would be interesting to examine the mean credibility levels of the same treatment when introduced as being a recommendation from family, friends or the media. Additionally, this adaptation of the CEQ seems psychometrically sound, and as such could easily be used by other researchers investigating different therapies. As such, one contribution of this instrument is to provide a more consistent and psychometrically validated framework of expectancy questions which other researchers can build upon. Such an instrument

is important given the large differences in expectancy measurement described in previous research. Such an approach would rely on the instrument being validated using different constructs and samples, however.

That being said, there are some important limitations to the study. Firstly, there was no behavioural or observational outcome to benchmark the test against. Secondly, although this study sampled twice from the student population and once from the staff population, this test still relies entirely on the responses of staff and students at one particular university. Further replication of the factor structure in heterogeneous groups is warranted before extensive use of the instrument takes place.

The major aim of this work was to develop and test an instrument which would capture more of the key factors surrounding treatment related expectancies. This aim has been achieved, in that the measure appears to be stable, and to correlate in the expected fashion with similar instruments. It does appear, from this work, that the conception of conventional treatments and alternative treatments appears to differ, and do not appear to form one general treatment credibility factor, but rather to have two distinct factors. This finding was used to inform the development of the IAT (see Chapter 6).

Chapter 6

Development of Implicit Measures

6.1 Introduction

The next stage in the research was to develop two IAT measures, one for treatment credibility and the other for optimism. The original approach taken in this study was as follows. Firstly, a number of interviews with doctors, complementary health practitioners and the general public (students) were carried out to assess the constructs associated with health. The results of this analysis are reported in Appendix B.

Following on from this, a repertory grid approach was employed to develop individualised stimuli for each IAT (described below in Section 6.2.1). This chapter describes that process, and then reports on the piloting of the IAT measures in a small sample, along with the piloting of the experimental research.

6.2 Methodology

6.2.1 Repertory Grids

The use of repertory grids in this research was as a bridge between the qualitative analysis carried out and the quantitative side of the research.

Repertory grids were developed by George Kelly as an aid to therapy (G. Kelly, 2003) although they have been used in many diverse situations in the ensuing years.

Repertory grids were developed out of Kelly's theory of cognitive consistency, an active area of research which fell out of favour following the discovery of cognitive dissonance by Festinger in 1947 (Greenwald et al., 2002).

The premise of the technique is simple. Firstly, participants are supplied with a list of important people in their life, such as their mother, an older sibling and a teacher whom they liked or disliked. They write down the names they have chosen for each person, and then they compare the people in groups of 3. For each group (or sort), they are asked to describe how two of them are similar and also how one of them is different in a word or short phrase. These words or phrases can then be analysed both quantitatively or qualitatively.

The approach taken in this project was as follows. In one of the surveys carried out on the UCC population (TCQ version 1) participants were asked to rank the most important people in their life who were related to health-care. This data was then sorted and ranked, and a list of the most common people used was compiled into a health related repertory grid. This was then administered to a small sample (N=17) to test the instrument. The results of this testing are described in Section 6.4.

6.2.2 Development of IAT

The plan for the development of the IAT was to use the constructs obtained from the repertory grids to develop useful stimuli for the IAT. Unfortunately, this portion of the research did not lead to a successful outcome, for reasons described below.

The Treatment Credibility IAT was developed from both the important figures which arose from the repertory grid, the qualitative interviews, and to match the explicit measure of treatment credibility.

The optimism IAT was developed along similar lines to most other Implicit Association tests, in that the survey for this measure was used as a base.

6.3 Development of Repertory Grid

The repertory grid approach was taken for the following reasons. Firstly, conventional methods of developing IAT measures use the published self-report scale as a base, but for placebo treatment expectancies, there is no real validated self report measure (the development of which was an aim of this study, reported in Chapter 5).

The first step was the development of stimuli for the repertory grid. This process was carried out along with the first administration of the TCQ, and utilised the same sample. Following the completion of the instrument, participants were asked the following:

Please rank, in order of importance, any people or qualities you believe are associated with health (use titles not names e.g. doctor, not doctor Murphy)

Table 6.1: Frequency of Exemplars, Rep Grid Study

Stimulus	Freq
doctor	209
nurse	142
pharmacist	48
physiotherapist	28
dentist	23
consultant	21
nurses	19
surgeon	19
doctors	16
exercise	11

As can be seen from Figure 6.1, and Table 6.1 the most popular exemplar for health was (unsurprisingly) doctor, followed by nurse, pharmacist and dentist. Additionally,

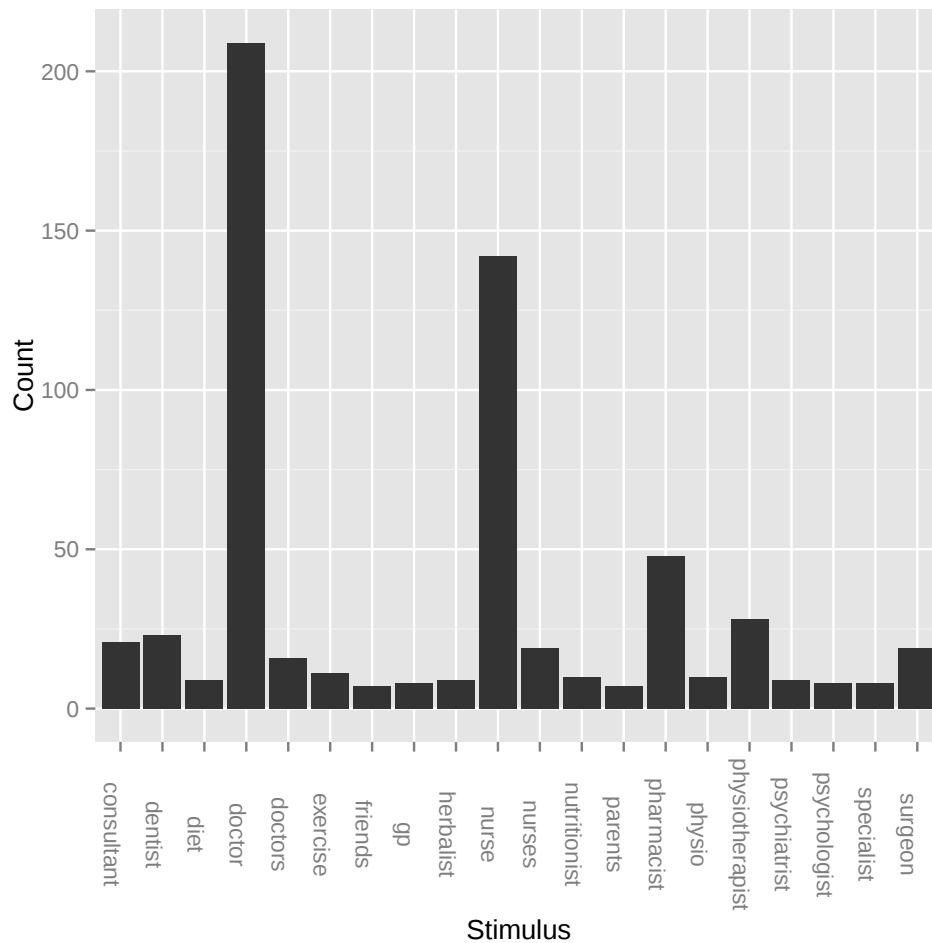


Figure 6.1: Histogram of Frequency of People or Qualities

a rep grid requires adjectives describing the exemplars, and these were taken from the standard repertory grid developed by Kelly.

6.3.1 Piloting of the Repertory Grid

This repertory grid (see Appendix C) was then piloted with a convenience sample. Unfortunately, it became clear that participants in this sample did not have enough examples of medical professionals in order to be able to accurately fill out the grid. The original pilot was carried out online (N=17) and not one respondent was able to complete the process. This was then followed with an attempt at card-sorting (N=3) where it again became clear that participants simply did not have enough experience with particular medical practitioners and environments, and so this attempt was abandoned.

6.4 Development of IAT

Therefore, the next step was to develop an IAT without the use of the repertory grid which had been part of the original plan. The approach taken here was two-fold, given the aims of the project. The first step was to use the interviews to establish a common set of themes, and then to match these to the explicit measure, in order to assure comparability. In the case of the Optimism IAT, this process was simple enough as there existed a self-report measure. However, in the case of the TCQ-IAT this approach presented some problems, as noted previously.

The eventual approach taken was to use forms of treatments for the stimuli, given that these were the closest to the explicit measure, and additionally because they had relatively high face-validity. The positive/negative stimuli were taken to be either words representing truth or falsity. This was done as it was believed that these words would be the most useful. The literature is relatively silent on the structured development of IAT measures. In general, self-report measures are used as a base, and the comparison dimension is typically self/other for many psychological constructs. The true/false distinction is novel in the literature, but makes sense for the purposes of this research.

The next step in this process was the piloting of the IAT and the examination of their associations with explicit measures of the construct, to ensure that they were measuring related areas.

6.4.1 IAT Stimuli

The stimuli used for each IAT were as follows:

6.4.1.0.1 Treatment Credibility IAT:

- Conventional: Creams, Pills, Surgery, Injections
- Alternative: Homeopathy, Acupuncture, Reiki, Flower Essence
- Real: Real, Accurate, True
- Fake: Fake, Inaccurate, False

6.4.1.0.2 Optimism IAT

- Optimism: Optimism, Happy, Improving, Succeeding
- Pessimism: Pessimism, Unhappy, Disimproving, Failing
- Self: Me, Mine, Myself

- Other: You, Yours, Yourselfes.

6.5 Piloting the IAT Measures

The IAT measures were piloted on 11 students from the psychology department. Each of these students were administered either the LOT-R and the optimism IAT or the TCQ and the TCQ-IAT.

6.5.1 Optimism IAT Pilot

The first step was to look at the composition of the sample. The sample was eleven students ($N=11$) of which ten ($N=10$) were female. The ages of the sample ranged from 19 to 27, with a mean age of 21.27 and a median age of 20. In this respect, the sample was quite typical of students in the psychology department. Students were asked had they heard of an IAT, to which all included in the sample replied no.

The mean Optimism IAT score (using the D algorithm, described in Chapter 7) was 0.5833, and the distribution can be seen in Figure 6.2. As can be seen from Figure 6.2, the majority of participants had a positive IAT score, representing their ability to respond quicker in the congruent rather than incongruent condition.

As can be seen in Figure 6.3, the correlation was not particularly strong, but there was a relationship there. Statistical examination of the correlation suggested that the relationship was equal to 0.15.

6.5.2 TCQ IAT

Next, the Treatment Credibility Questionnaire was examined to assess its correlations with the self report measure, in preparation for the pilot.

The range is greater for the TCQ IAT, potentially because of the less familiar nature of the stimuli. Unfortunately, the demographics were not available for this pilot, due to data loss, and so this is not reported on explicitly. However, from a report prepared at the time, the correlation was equal to 0.20, which is in line with that for the optimism IAT.

6.6 Piloting the Experimental Procedure

The participants in the pilot study ($N=7$), consisted of five men, one woman and one participant who did not report gender. The mean age of the participants was 25.3 ($SD=4.13$).

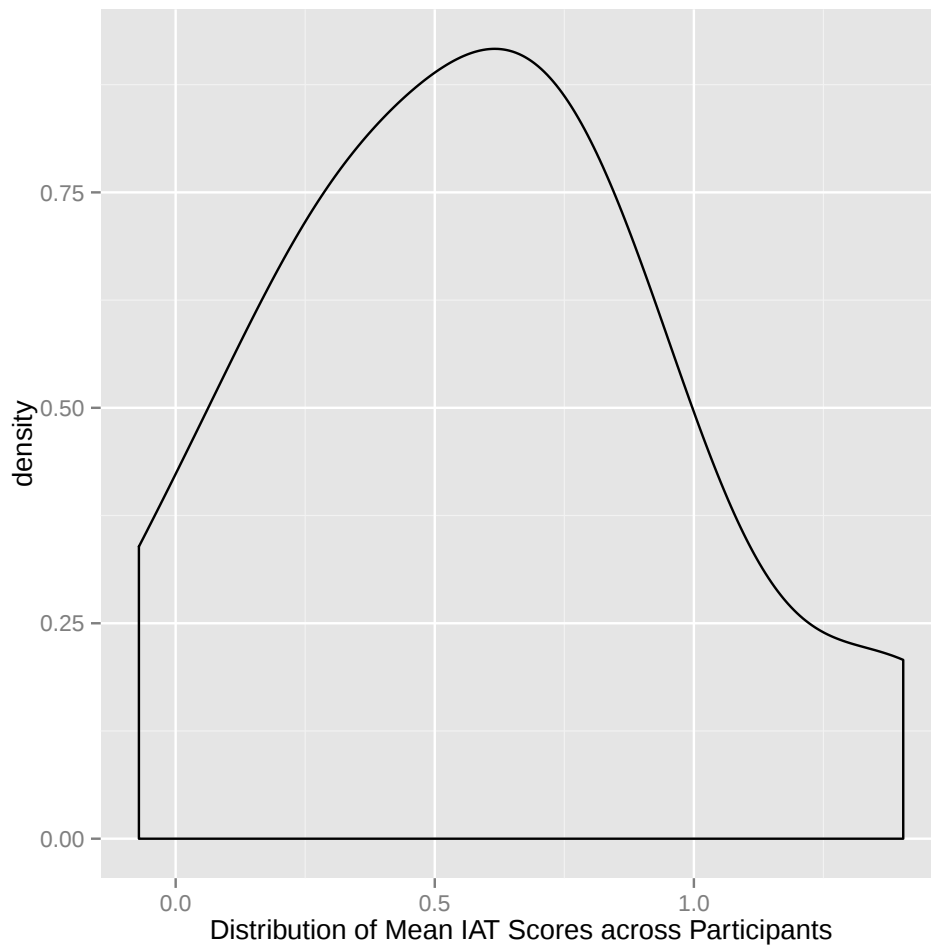


Figure 6.2: Distribution of Optimism IAT Scores

Table 6.2: Gender of Pilot Participants

	Gender
Female	1
Male	5
Not Supplied	1

Figure 6.4 shows the correlations between the various explicit measures in this study. It can be seen that the negative correlation reported on between the LOT-R and the MAAS replicated in this sample, suggesting that it was not a statistical fluke. It can also be seen that the conventional and alternative treatment credibility items correlated with one another, but not really between Conventional and Alternataive methods (apart from the strong negative correlation between Injection and Acupuncture credibility, which also occurred in Study One with the TCQ (Chapter 5). However, with such a small sample size, these results should be regarded as a little suspect.

A Kruskal test on the differences between blocks and total time gave a Chi Square of

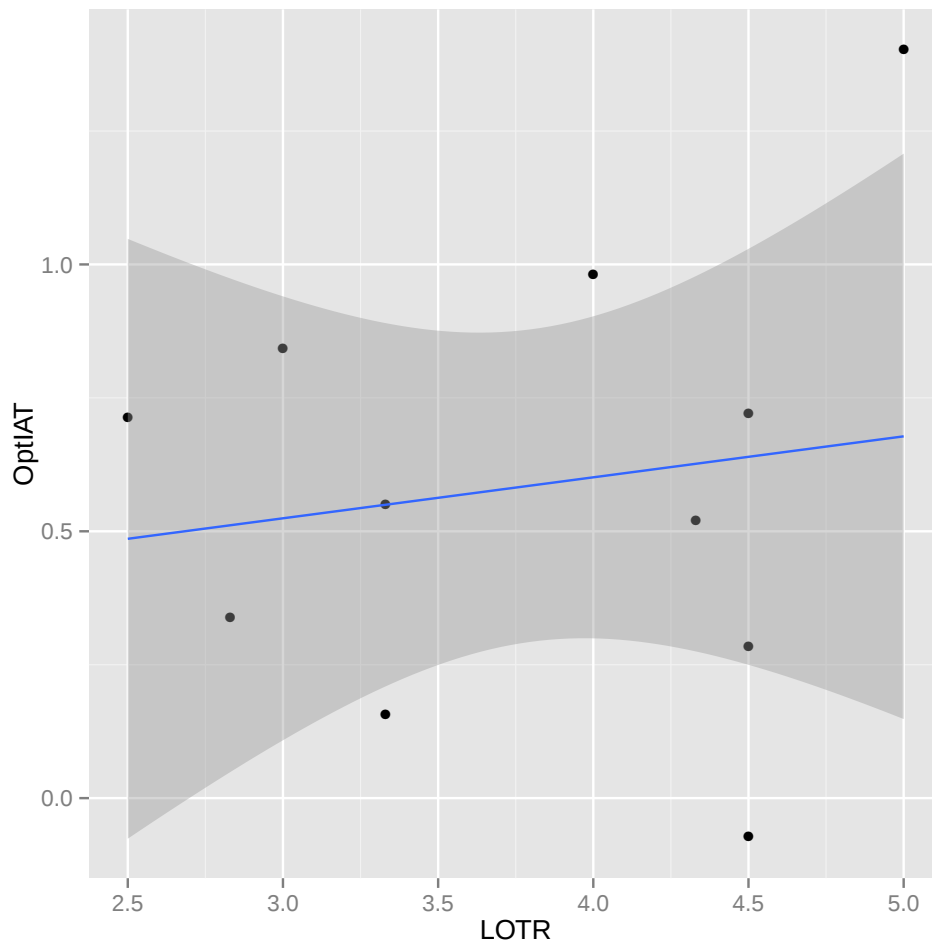


Figure 6.3: Scatter-plot of Correlations between LOTR and Optimism IAT

23.59, with an associated p value of less than 0.0001 with degrees of freedom equal to 4. This result is somewhat obvious, as blocks three and five are combined tasks, and would be expected to show a different pattern. This is in line with expectations before this experiment.

The first step was examining the correlation between the two explicit measures. This was found to be 0.62, using Spearman's ρ . This is somewhat higher than would be hoped for, but still only represents 36% of the variance in common between the two measures. It is possible that this is due to the small sample, and as the previous piloting studies did not show this pattern, this explanation seems to be the most likely.

The Spearman correlations with the MAAS were 0.37 (for the Treatment Credibility IAT) and 0.08 (for the Optimism IAT).

Surprisingly (based on the results of the pilot described above), the correlation between the optimism IAT and the LOT-R was negative, but was of the same magnitude as was seen in the previous study ($\rho = -0.26$).

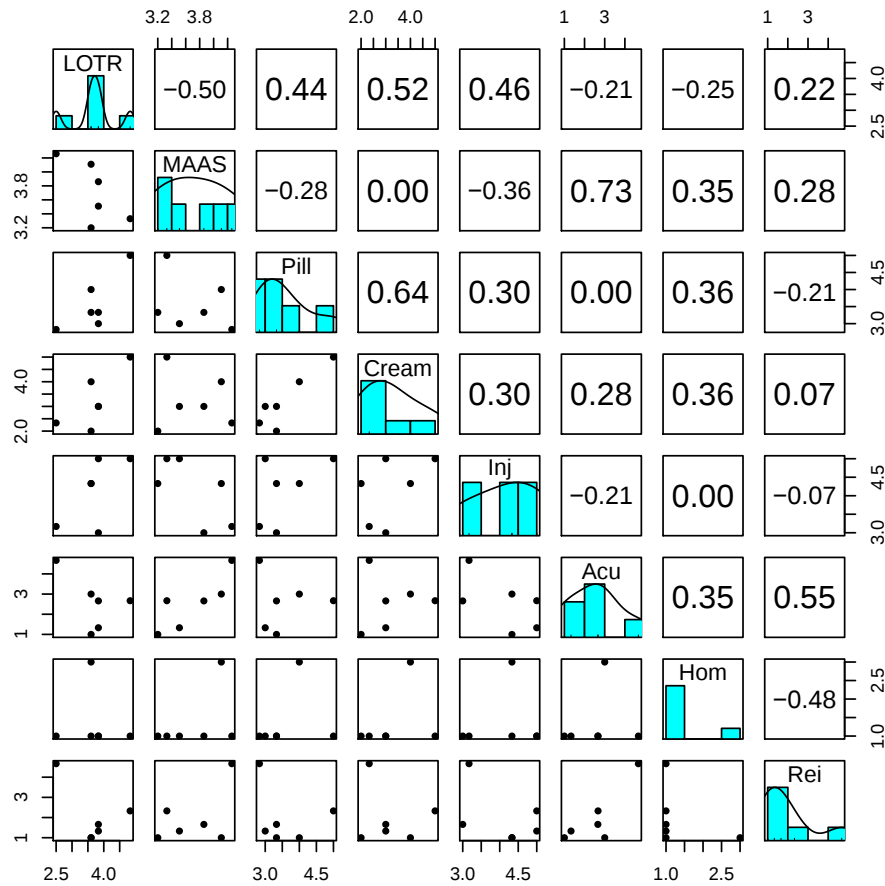


Figure 6.4: Pairs Plot of Explicit Variables, Pilot Study. LOTR: Life Orientation Test, Revised. MAAS - Mindful Attention Awareness Scale. Pill - Pill Credibility. Cream - Cream Credibility. Inj - Injection Credibility. Acu - Acupuncture Credibility. Hom - Homeopathy Credibility. Rei - Reiki Credibility

Figure 6.5 shows the pain ratings of all participants over time. There is definitely a upwards trend in the pain measurements, as would be expected given the design of the experiment, where pain scores increase over time. There is substantial variability in the responses, as can be seen from the wildly different lengths of the series. Two participants in the no treatment condition went for longer than 40 minutes, while none of the participants in the treatment condition did so. This is obviously a problem, given the goal of the experiment, and the results are illustrated graphically in the preceding figure.

Table 6.3: Predicting Placebo Response from TCQ IAT Scores

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-0.4296	1.0351	-0.41	0.6782
TreatmentIAT	-0.0019	0.0026	-0.73	0.4657

As can be seen from Table 6.3 and 6.4 the relationship between the outcome variables

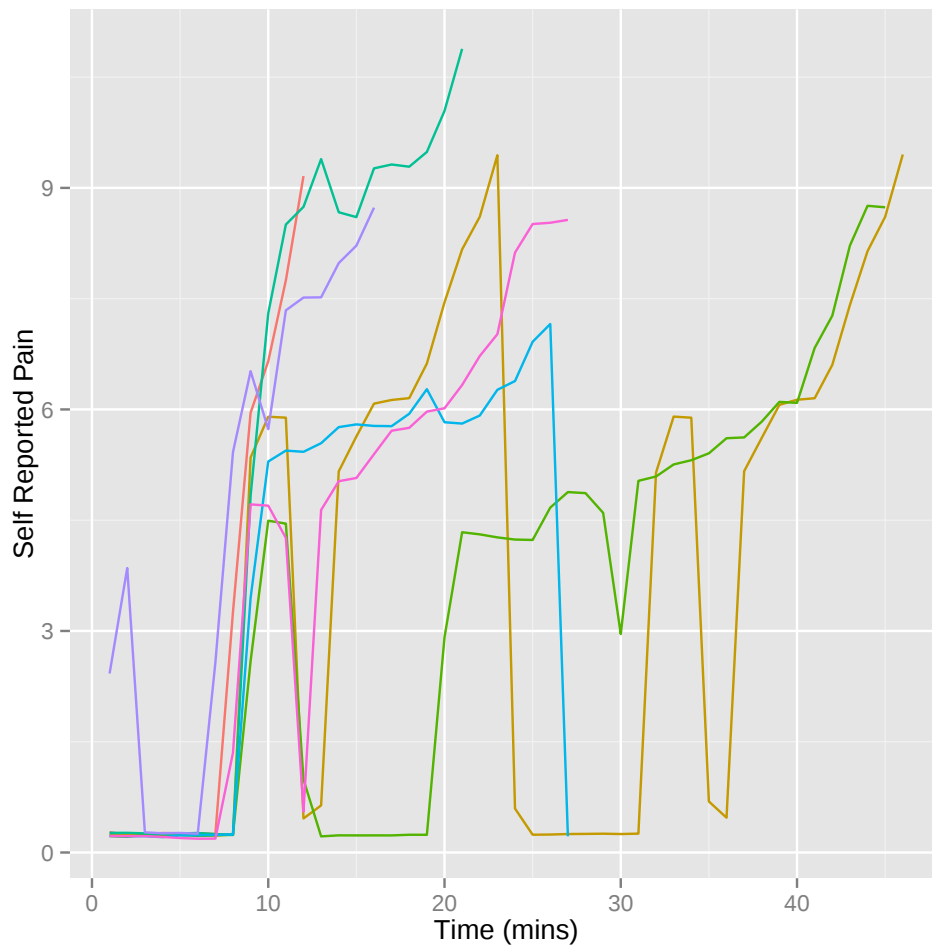


Figure 6.5: Participant individual pain responses over time

Table 6.4: Predicting Placebo Response from OPT IAT Scores

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-4.0792	6.3541	-0.64	0.5209
OptimismIAT	0.0035	0.0055	0.65	0.5185

Table 6.5: Predicting Placebo Response from Explicit Credibility Scores

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-8.1988	9.5410	-0.86	0.3902
LOTR	2.4673	2.6614	0.93	0.3539

did not appear to be large, though they were all (with the exception of the Treatment Credibility IAT) in the correct direction. Note that the generalised linear model for the ConvAltRatio did not converge, and the coefficients are not reported.

Additionally, only 3 of the seven participants responded to placebo, suggesting that a larger sample size would be required to show a significant effect. This problem was worked around by using a scrambled sentence task on the majority of placebo participants, along the lines of Geers (Geers, Helfer, et al., 2005).

6.7 Discussion

This chapter consisted of three smaller parts. Firstly, the results of the repertory grid approach to development of the IAT's was reported, and unfortunately this approach did not prove successful. It may be that this method may prove useful for different IAT's, but it did not prove useful here.

Secondly, the IAT measures were developed and piloted in small samples. This process found that each measure correlated moderately with the appropriate explicit measure, which is reassuring for the goal of the experiment.

Finally, the experimental procedure was piloted to assess what the appropriate sample size should be, and if further changes should be made to the setup.

A number of issues became clear as a result of this process. The first is that the effect sizes achieved from placebo suggestions were quite small. The solution to this in the larger study was to include a priming manipulation (the scrambled sentence task) for the majority of participants, a small group not being primed to ensure that the effect of this manipulation could be tested.

Most importantly, there appears to be a (small) relationship between the explicit and implicit measures, which would seem to indicate that these measures can be effectively used in the experimental setup.

Additionally, the use of the priming manipulation allows for the major hypotheses of the study (the relationship between placebo response, explicit and implicit measures) to be tested more efficiently.

Additionally, another change which was made to the study was the introduction of a placebo condition, whereby participants in this condition were told that they would receive a placebo. This change was on foot of the study by Kirsch *et al* whereby such a manipulation was found to be effective in a placebo/no treatment design. Such an effect would clearly be of interest in placebo analgesia, and this would add to the usefulness of the study. This also enabled the larger experimental study to assess the impact of both strong (Deceptive Administration) and weak (Placebo) suggestions, which is important if either of these types of suggestions are predicted by different sets of explicit and implicit measures.

Chapter 7

Measuring Placebo by Multiple Methods

7.1 Introduction

This chapter describes the experimental portion of the research, which involved the assessment of treatment expectancies using implicit and explicit measures. This section further involved the testing of a number of proposed theoretical models around the relationship of explicit, implicit and physiological measures to the placebo response. The justification behind these theoretical models is described in Chapter 3.

The placebo effect is a complex phenomenon. It has been established that it can be predicted by some variables which are typically measured using self report approaches (Kirsch, 1997; Pollo & Benedetti, 2009; Whalley et al., 2008; Geers, Helfer, et al., 2005). These variables include expectancies and optimism. Optimism was measured using the Life Orientation Test, Revised (LOT-R - see Chapter 4 for details on this instrument), and a new measure for treatment credibility expectancies was developed in Chapter 5 (the Treatment Credibility Questionnaire) to assess treatment-related expectancies.

Explicit measures only appear to explain a small proportion of the variance in the observed placebo response (given a correlation of 0.3 with the response, this equates to less than 10% of the variance). It is the contention of this thesis that some of the residual variance can be predicted by the use of implicit measures (specifically the IAT). Therefore, two implicit measures, one of Treatment Credibility and the other of Optimism were developed (see Chapter 6 for details).

The relationship between the observed placebo response, and these explicit and implicit measures was of primary importance to this research, and so a measure of Mindfulness (Mindful Attention Awareness Scale) was also used. The rationale behind this choice of measure was explained in Chapters 2 and 4, but briefly, this measure has been shown both to correlate with a cluster of other variables around cognition and implicit attitudes (Levesque & Brown, 2007; M. T. Conner et al., 2007). Additionally, this variable is also related to self-reported health (see Chapter 4 for further details), and as such is a potential moderator of the placebo response itself.

In order to make use of psychometric modelling, large samples of the self report instruments were collected in the same population from which the experimental participants were drawn. This allowed for factor score and IRT models to be built for each of the measures. This work was reported in Chapters 4 and 5.

Physiological (GSR or skin conductance) recordings were also collected in order to examine the physiological characteristics of placebo and to increase the validity of the study in the minds of the participants.

The collection of these various forms of data, along with a behavioural criterion (the observed placebo response) allowed for psychometric models (Structural Equation

Models or SEM) to be examined to separate out the effects of the various predictor types.

In essence, this thesis (and more specifically this chapter) seeks to marry the strengths of psychology in psychometric modelling to its complementary strengths in experimental design, with the aim of establishing these methods and measures relative usefulness in the prediction of placebo.

Along with testing the major hypotheses, SEM models were applied to examine some proposed theoretical structures. Much more detailed descriptions of the major models applied are given in Chapter 3, but they are reviewed below.

7.1.1 Theoretical Models

- The direct expectancy model of Kirsch, in which the effects of all other variables are mediated through expectancies (both implicit and explicit)
- The conscious expectancy model of Stewart-Williams *et al* whereby implicit measures have no effect
- A model whereby optimism mediates the effects of all other variables on placebo response
- Two models which allow only implicit or explicit expectancy measures to be the mediators of the placebo response.

This chapter proceeds with the following sections:

1. A short rationale for this study
2. A description of the methodology employed in this chapter to test the major hypotheses
3. A testing of the major hypotheses around the placebo and its relationship to implicit and explicit treatment expectancies.
4. A discussion outlining the major issues raised by this chapter.

7.1.2 Placebo Effects and Implicit Measures

As discussed in Chapter 2, Section 2.12.1, there are some reasons to believe that the placebo effect is impacted by implicit influences. These are the following:

- Research shows that priming can increase the size of placebo effects;
- The results of Kirsch (1998) which showed that informed conditioning was less effective than uninformed conditioning would suggest that the effectiveness of

placebos can be reduced by conscious processing;

- It has been suggested that the IAT captures substantial state variance (as evidenced by low test-retest reliability) (Egloff et al., 2005) and it has also been argued that response to placebo is similar (Whalley et al., 2008), thus suggesting that there may be a link;
- Semantic priming has been shown to exert changes without becoming apparent to consciousness again suggesting that an implicit measure might prove useful in measuring this effect.

The implicit measure chosen for this purpose was the Implicit Association Test (IAT). This choice was made both because it is the most widespread, and as such more of the best practices and problems with the measure were available in prior research.

These problems (discussed at length in Chapter 2, Section 2.8), have included:

- Low test-retest reliability (approximately 0.49)
- method variance affecting the results
- Issues around the scoring and interpretation of the results

This research addresses these problems in the following manner. Firstly, the low test-retest reliability is an issue, and it is also possible that any effect would be contaminated by attitude-behaviour consistency.

However, the one counterpoint may be that IATs' are more useful in predicting over short time scales, as could be argued from the findings that they predict spontaneous behaviour better than explicit measures (Richetin et al., 2007; Steffens & König, 2006; Steffens, Kirschbaum, & Glados, 2008))

The method variance problem was moderated by the use of the D measure from Greenwald (Greenwald, Nosek, & Banaji, 2003), as demonstrated by Mierke (Klauer & Mierke, 2005). The impact of these potential confounders was examined by assessing the correlations between the two IAT measures used in this study (Treatment Credibility and Optimism) and if these were too high ($r \geq 0.3$) then controlling for them by including both IAT's in each regression model.

The issues around the scoring of the results (connected to the use of the mean with non-symmetrical distributions (Venables & Ripley, 2002)) were examined by scoring the IATs using both the mean and median as measures of central tendency.

7.1.3 Placebo effects and Deception

The commonly accepted model of the placebo effect claims that deception is necessary for the effect to occur. The entire theory behind randomised controlled trials suggests

that because participants respond to the mere administration of medicines, then a control (the placebo) is needed to allow the true effect of the drug to be established. A necessary part of this design is that the participants in the placebo group need to believe that there is a chance that they are getting the real treatment.

Some theorists have argued against this (van Deventer, 2008; Evans, 2002), and a recent randomised controlled trial has shown that compared to no treatment, an open placebo can perform significantly better (T. Kaptchuk et al., 2010).

The Kaptchuk *et al* study, which was a three week randomised controlled trial using IBS patients showed that administration of an open-label placebo was associated with a significant improvement ($d=0.79$) at in the IBS Global Improvement Scale (the major outcome measure) the 21 day endpoint for the study. Similar trends were observed for other symptom severity ($d = 0.53$) and quality of life ($d = 0.40$)

However, this study did not include a deceptive placebo condition, and so is not a true test of the theory that deception is unnecessary for the placebo response. One secondary aim of this study was to examine the theory that placebo responses can be induced without deception.

7.1.4 Explicit Measures and Placebo

A number of issues are apparent in the current research around treatment-related expectancies. The first issue is that expectancies are typically measured with a simple one question scale. One of the aims of this thesis was to apply psychometric methods to the development of a more sophisticated measure of treatment credibility and expectancies (see Chapter 5). The other problem is that expectancies are assumed to be conscious, even though they appear to have far more in common with unconscious responses than they do with controlled processes.

The measure developed in Chapter 5 attempts to address the first problem, while the IAT's developed in Chapter 6 address the second. The measure adapted in Chapter 5 had six questions for each of the six different treatment modalities, which included conventional (Pills, Creams, Injections) and Alternative (Acupuncture, Homeopathy, Reiki) treatments. Three of the questions related to expectancy, while three related to credibility (factors found in the original development of the instrument). These 36 questions were reduced to 18 using factor analytic and IAT measures, to allow for easier administration (c.f. Section 5.11).

Additionally, this experiment also investigates the relationship between expectancies and optimism, given that optimism is often defined as “generalised outcome expectancies” there would seem to be a relationship, but whether both of these constructs capture different facets of the response or merely co-vary with one another

is not clear, and this was investigated using regression and structural equation models. The relationship between these two variables is discussed further in the section on theoretical models in Chapter 3.

7.1.5 Aims and Objectives

The aims and objectives of this portion of the research were as follows:

- Examine the relationship between the two implicit measures
- Examine the relationship between the explicit and implicit measures
- Assess the similarities between the survey samples and the experimental sample on the explicit measures
- Examine the relationship between implicit and explicit measures of treatment credibility and optimism and the placebo response;
- Test a number of theoretical models to describe this relationship;
- Examine the relationship between placebo response, pain levels and the physiological variables.

7.2 Methodology

7.2.1 Experimental Procedure

7.2.1.1 Recruitment

Recruitment for participants primarily took place via email. In total, six emails were sent to random sub-sets of the student mailing list from the period January 2nd until April 11th. Given that the experiment took place in the Applied Psychology department, somewhat off campus and that the experiment involved suffering painful stimuli, an inducement of a smartphone was offered to one participant who completed the procedure on the basis of a draw following the completion of the research. The experiment ran from January 17th until April 14th 2011 inclusive.

7.2.1.2 Measures

The following measures were used in this experiment. Firstly, age, gender and course of study were collected for each participant. The MAAS and LOTR were also administered to each participant, as was a shortened version of the treatment credibility questionnaire (described in Chapter 5, Section 5.11). Following the

completion of the explicit measures, participants completed a Treatment Credibility IAT and an Optimism IAT. The stimuli used in each of the IAT's were as reported in Section 6.4.1, Chapter 6, where the piloting of this instrument is described.

In addition, participants gave verbal reports of their pain levels to the experimenter at one minute intervals, and these were recorded (along with condition and exact time of application of bandage and when the squeezing stopped) on a sheet of paper, along with the participants identification number.

7.2.1.3 Procedure

All participants (N=111) were met at the entrance to the building by the primary researcher. They were given the informed consent documentation, and after they signed it, they completed three questionnaires (the MAAS, the LOT-R and the TCQ). Following this, they completed both an Optimism IAT and the Treatment Credibility IAT, where order of administration was counterbalanced across participants.

Following this, the participants sat down next to the Biopac physiological monitoring data, and baseline data was recorded for five minutes. Then, a blood pressure gauge was wrapped around the upper part of the non-dominant hand of the participant, and they were asked to squeeze a hand exerciser twenty times for two seconds each time. This method is known as the sub-maximal tourniquet method. One minute after this, and every minute thereafter, participants were asked to rate their pain on a VAS from 0 to 10.

If the participant was in the Deceptive or Open Placebo group, then when they rated their pain as 7 or higher, the placebo cream was applied. The experiment continued until the participant either decided to withdraw, their pain rating reached 10 or 45 minutes elapsed from when the bandage was applied. EDA recordings were taken one thousand times per second using the Biopac equipment and VAS ratings were recorded on paper by the experimenter. The placebo cream consisted of moisturiser in a pharmaceutical container, and was unlabelled.

Participants in the treatment group were told that *the cream was a potent painkiller, recently approved and proven to reduce pain, which would take effect almost immediately*. Participants in the placebo group were told that *they were receiving a placebo and that placebos have been clinically proven to reduce pain, and that it would take effect almost immediately*.

7.2.2 Analysis of Experimental Data

The first step in the analysis of experimental data was to examine the comparability of each of the different groups. This procedure was carried out using analysis of

variance techniques. There were no differences in pre-treatment levels except for GSR levels, which were different at baseline between the Deceptive and Open Placebo groups. This difference was not significant, though it was visible.

7.2.2.1 Analysis of Reaction Time Data

Reaction time data has been studied by (mostly cognitive) psychologists for many years. The Implicit Association Test has been used in almost 300 published papers and reports. However, with a few exceptions, there has been almost no overflow from one area of study to the other.

The typical approach to analysis of IAT data goes as follows (Greenwald et al., 1998): firstly, the data is checked for outliers. Outliers, in this case are defined as responses less than 300ms and greater than 3000ms. Any such outliers are recoded to 300 or 3000ms respectively. Following this procedure, a mean is taken of the items in each condition. These means along with standard deviations are reported. The response latencies are then log transformed (to reduce positive skew) and the IAT score is calculated as follows.

Given the participants mean latency in each condition, their IAT score is the mean for the incompatible condition (i.e. White + Unpleasant) less the mean from the compatible condition (i.e. White + Pleasant) divided by the average of the two within group standard deviations. This measure is typically called D , and was developed after analysis of an extremely large sample of IAT responses (A. G. Greenwald et al., 2003). In addition to the change of scoring procedure in the 2003 revisions, the threshold for outliers was also substantially widened to 10000 ms. Given that the sample used for this re-analysis was over one hundred thousand, this widening of the threshold was presumably based on experience with a much broader population than was used in many early IAT experiments.

While means are useful summary statistics, they are most optimal in situations where the distribution is unimodal and symmetric (Venables & Ripley, 2002). Reaction latency data are neither, so the choice of mean seems to have been made from familiarity rather than principle.

Given the typical right skew observed in reaction time distributions, the median would seem to be a much better measure of central location than would the mean. This right skew typically occurs as there is a hard bound on how quickly a participant can respond, but no such bound (unless enforced by the procedure) in the maximum time taken to respond. Therefore, in this research, the efficacy of these different measures of central tendency will be reported.

The differences between the mean and median based approaches were examined and reported, to determine if it makes any difference. In this experiment the mean and

median IAT scores were highly correlated and the use of one measure versus the other made no difference to the results.

Manipulation checks during the course of the experiment revealed that no participants were familiar with the IAT.

7.2.2.2 Analysis of Implicit and Explicit Measures

Next, the relationships between the explicit measures individually, and the implicit and explicit measures collectively, was assessed. This was done by examining their inter-correlations, and also by comparing them to the results of the earlier survey samples. Finally, the factor score models from the earlier studies were applied to the experimental data, to create new synthetic variables to test the efficacy of this approach.

The cross-correlations of the pain ratings with the GSR readings were examined as part of the placebo analgesia experiment, and the breakdown of pain and GSR by condition, placebo response and gender was examined.

The pain data was also used to classify participants as either placebo responders or non-responders. This was operationalised by examining if their pain levels decreased following administration of placebo. If this happened, they were classified as placebo responders, and if not, they were classified as non-responders.

The major hypotheses of this study were then tested using a ten-fold cross-validation procedure and stepwise logistic regression. Note that as discussed earlier, the use of a test-set to provide p-values allows for accurate inferences to be made.

In addition, these logistic regression models were compared against models using both factor scores and IRT ability estimates as predictors to determine the usefulness of this model based approach.

7.2.2.2.1 Analysis of Physiological Data Skin conductance recordings (GSR) were taken as part of this experiment. This source provided recordings at 1000Hz for the entire procedure.

7.2.2.3 Structural Equation Models

Finally, the relationships between physiological responses and the psychological (both implicit and explicit) variables collected, were examined. More details on the hypotheses regarding this can be found in Chapter 3. Briefly, the models tested were as follows:

- My model, whereby implicit and explicit expectancies jointly mediate all other variables
- A model like that of Kirsch, in which expectancies mediate all other variables
- A model where all effects are mediated by Optimism
- Model(s) where either implicit or explicit expectancies are solely responsible for the observed effects
- A null model

These models allowed for the principled comparison of a number of different theoretical models on the relationship between expectancies, optimism and physiological variables on the placebo response.

7.3 Results

7.3.1 Data Validation

In all, 114 responses were collected for the explicit measures. Following dropout, 113 responses were collected for the IAT measures. Following a fuller description of the experiment, but after administration of the explicit and implicit measures, 3 participants did not proceed any further (at their own request, in line with APA guidelines). Therefore there were 110 participants left for further analysis.

However, as part of the experimental design, participants were asked to pick a four digit code to represent their data across the the different parts of the experiment (explicit measures, implicit measures, physiological data and pain ratings). This was done to prevent the researcher from inferring participant data based on the order of administration to each participant. Unfortunately, this meant that a number of participants picked the same number. This was noticed during the data entry phase of the research and some of the problems could be corrected based on date and time information. However, given that this information could not be generalised across the physiological datasets without error, it was decided to remove all participants who had a duplicated participant number. This approach was chosen as it is the most conservative and the least likely to cause inferential error (as duplicated data will impact means and tail probabilities more than the removal of data). After this process, 100 participants remained, and this was the final sample used in all analysis below.

7.3.2 Randomisation Checks

Firstly, the comparability of the groups were assessed to ensure that the randomisation process had proved effective.

Table 7.1: Summary of ANOVA of Treatment Credibility IAT Scores by Condition

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Condition	2	0.13	0.06	0.12	0.8911
Residuals	100	54.74	0.55		

Table 7.2: Summary of ANOVA of Optimism IAT scores by Condition

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Condition	2	1.02	0.51	1.06	0.3492
Residuals	100	47.76	0.48		

As can be seen from Table 7.1 and Table 7.2 the scores on the treatment credibility IAT or the optimism IAT did not differ by Condition.

The results of a chi-square test showed that the gender of participants in each condition were equivalent, ($p=0.633$). In addition, the priming manipulation was not significantly different across groups.

The counterbalancing was checked for influence on the IAT scores, and no significant differences were found, suggesting that order of administration of the IAT's had no impact on the results.

However, there was a difference in the mean (and median) GSR scores at baseline. Baseline measurements were taken for the first one hundred seconds of the experiment, and can be seen in Table 7.3 broken down by condition, though from an examination of the standard deviations, these differences do not appear to be significant. The reasons for this difference are unclear, given that no participant was informed of their condition allocation until after their self-reported pain levels reached seven.

Table 7.3: Mean and SD of GSR by Condition at baseline (first 100 seconds)

Condition	Mean	SD
No Treatment	7.92	6.70
Placebo	9.58	7.33
Treatment	11.71	8.11

7.3.3 Analysis of IAT data

The first step in the analysis of IAT data is to examine the differential impact of using the mean versus the median as the measure of central tendency for the calculation of IAT scores (the D measure). The results showed that there were no major changes attributable to this difference (the correlation between the two scores was $r = 0.91$).

Next, we examine the difference between the mean and median scores for the TCQ IAT. There was little difference between the two measures of central tendency ($r = 0.89$).

The next question is whether or not the IATs have been contaminated by method variance. This can be assessed by examining the correlations between the Treatment Credibility and Optimism IAT. The correlation between the two mean scored IAT measures was ($r = 0.003$), while the correlation between the two median scored IAT measures was $r = 0.1$, thus showing that method variance does not appear to have contaminated the results.

An analysis showed that the majority of items were responded to relatively quickly in both categories. Of note, however, are the outliers which were words which were associated with fake treatments in Block 3 (where they were paired with conventional treatments) and words which were associated with real treatments (where they were paired with alternative treatments). This would seem to suggest that the words were in fact serving their intended purpose.

A similar pattern emerged from the conventional and alternative stimuli. Interestingly, it appears that response times were slower overall in Block 5, which may represent fatigue. However, the order of IAT's was counterbalanced, so one would expect to see the same pattern in the Optimism IAT's if this was the case.

The same pattern emerged for the optimism IAT in that the Block 5 scores were much more variable and overall participants responded slower to this block, suggesting that this did occur as a result of fatigue. The pattern of Block 5 responses tending to be slower was repeated in the mean response latencies for the positive and negative words in the Optimism IAT. Note that Disimproving appears to be the word with the highest mean latency, which is not surprising given its relative unfamiliarity (compared to the other words, at least).

Next, the correlation between the different block times is assessed.

Table 7.4 gives the exact Kendalls τ between each of the blocks. As can be seen the correlations hover between 0.3 and 0.4, which is in line with expectations prior to the experiment.

Table 7.4: Correlations between the blocks of the treatment credibility IAT (Kendalls tau.) All correlations are significant at the $p < 0.001$ level

	TCQBlock1	TCQBlock2	TCQBlock3	TCQBlock4	TCQBlock5
TCQBlock1	1.00	0.45	0.33	0.30	0.37
TCQBlock2	0.45	1.00	0.35	0.21	0.37
TCQBlock3	0.33	0.35	1.00	0.33	0.26
TCQBlock4	0.30	0.21	0.33	1.00	0.31
TCQBlock5	0.37	0.37	0.26	0.31	1.00

Next, the same process is repeated for the Optimism IAT.

Table 7.5 shows that the correlations are a little higher than for the Treatment Credibility IAT, but still within an acceptable range. Its interesting to note that (with the exception of Block 5), the correlations are strongest between adjacent blocks, and drop off as the blocks move further apart, suggesting that there may be correlations between adjacent blocks (which would not be surprising given the nature of the IAT task).

Table 7.5: Correlations between the blocks of the Optimism IAT (Kendalls tau). All correlations are significant at the $p < 0.001$ level

	Block 1	Block 2	Block 3	Block 4	Block 5
Block 1	1.00	0.46	0.33	0.31	0.37
Block 2	0.46	1.00	0.46	0.43	0.41
Block 3	0.33	0.46	1.00	0.44	0.35
Block 4	0.31	0.43	0.44	1.00	0.42
Block 5	0.37	0.41	0.35	0.42	1.00

There were no significant differences ($t = -0.4973, p = 0.6211$) between men and women in the sample with regards to their scores on the Treatment Credibility Questionnaire. However, the variance was much higher for men, which was a pattern replicated in previous research into Treatment Credibility (using a self report instrument described in Chapter 5). The same pattern was repeated for the Optimism IAT, in that there were no significant differences ($t = -0.8234, df = 49.761, p = 0.4142$) between males and females in this sample.

7.3.4 Explicit Measures

Optimism and mindfulness were positively correlated with one another (see Figure 7.4, on page 170). Additionally this correlation appeared to be relatively stable across condition, though it appeared a little weaker in the Deceptive Placebo Group. This is in contrast to the results found in a much larger scale study carried out earlier in the research. Note that one plausible explanation for this effect is that, in the experiment, the measures were administered in the opposite order - Optimism, followed by Mindfulness. It is possible that the completion of the mindfulness measure affected the way in which participants approached the Optimism measure. This theory is more fully discussed in Chapters 4 and 8.

7.3.5 Relationships between Experimental Samples and Survey Samples

Given the focus of this thesis on the integration of survey and experimental research, the next step was to examine the differences and similarities between the samples

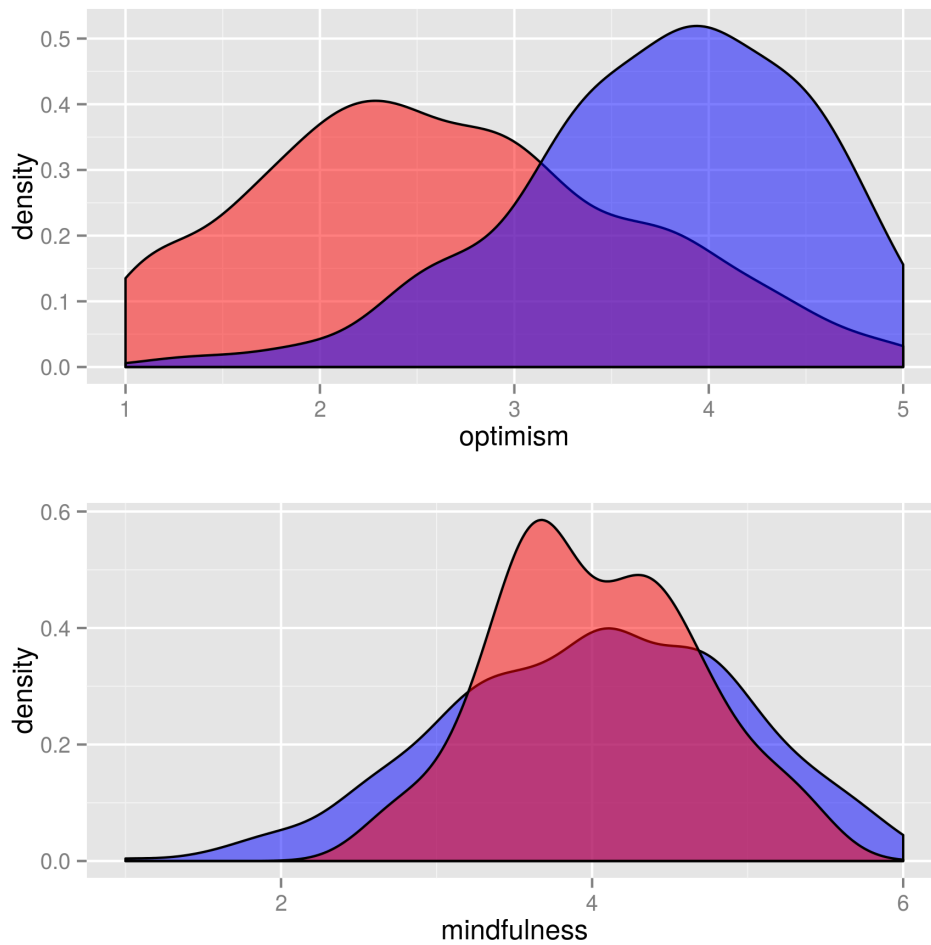


Figure 7.1: Density Plot for Optimism (top) and Mindfulness Scores (bottom), Survey Sample (red), and Experimental Sample (blue)

collected from the general population via survey and the experimental sample.

From Figure 7.1 (top panel) it can be seen that the two distributions are extremely different, with a much higher average optimism score in the experimental sample. To some extent, this is not unexpected given that the study was described as an investigation of painkilling drugs and there was an opportunity to win a smart-phone, so perhaps students with higher levels of optimism were more likely to agree to participate.

As can be seen from Figure 7.1 (bottom panel), the pattern was quite different for mindfulness levels (as measured by the MAAS) as the levels of mindfulness were less long-tailed and more concentrated around a central peak in the survey sample. Again, this may be due to the association of mindfulness with introversion, as introverts may have been less likely to respond to the email invitation(s) to take part in the study, while an online study might not have suffered from the same problem.

Finally, the treatment credibility questionnaire scores were examined to assess the

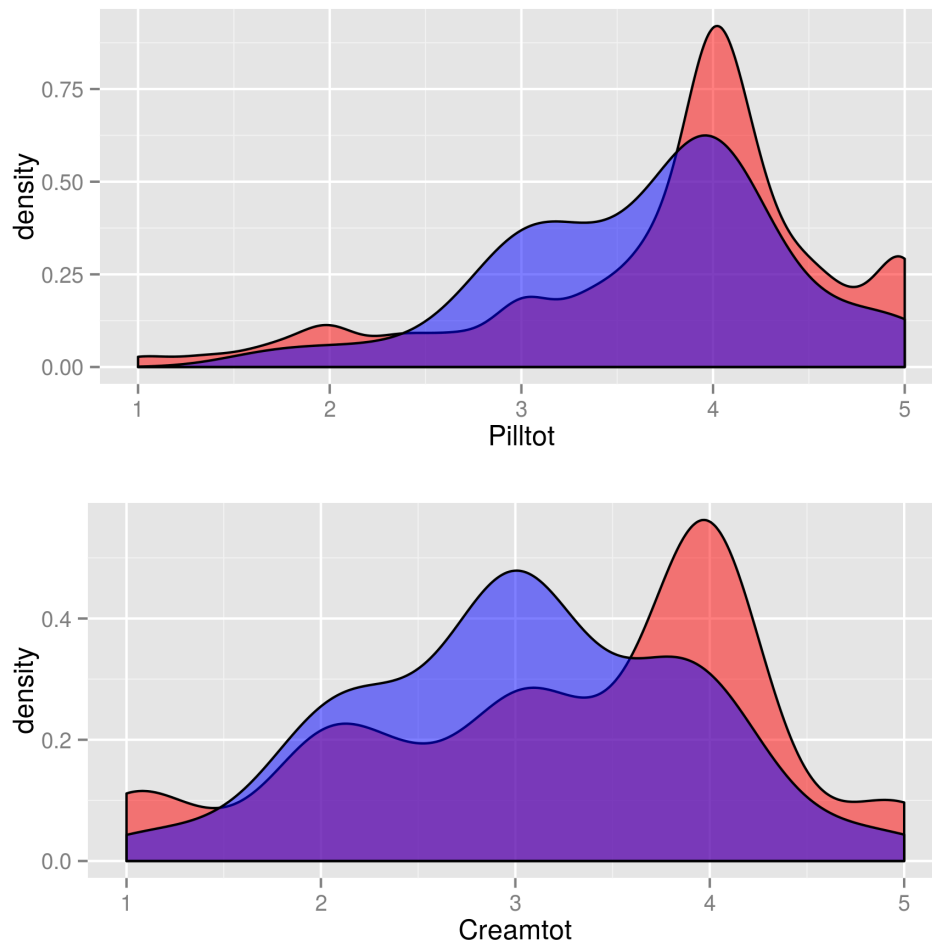


Figure 7.2: Density Plot for Pill Credibility (top) and Cream Credibility (bottom), Survey Sample (red), and Experimental Sample (blue)

differences between the survey and experimental samples.

First, the differences between the two samples in terms of Pill credibility were examined.

As can be seen from Figure 7.2 (top panel), the general population sample was higher peaked, with less variation around the peak than was the experimental sample. In fact, the experimental sample seemed to be more variable than the survey sample, which could either be due to a true difference in the distributions or due to a greater uncertainty in the experimental sample due to the smaller sample size.

As shown in Figure 7.2 (bottom panel), the survey group tended to have a more positive view of painkilling creams. While the survey group is strongly peaked at the right of the plot, the experimental group were more evenly distributed, with a peak at the centre of the plot. This is interesting, as one might expect the experimental group to be more positive towards painkilling treatments in general, given that they had agreed to take part in a study which purported to examine the effects of a new

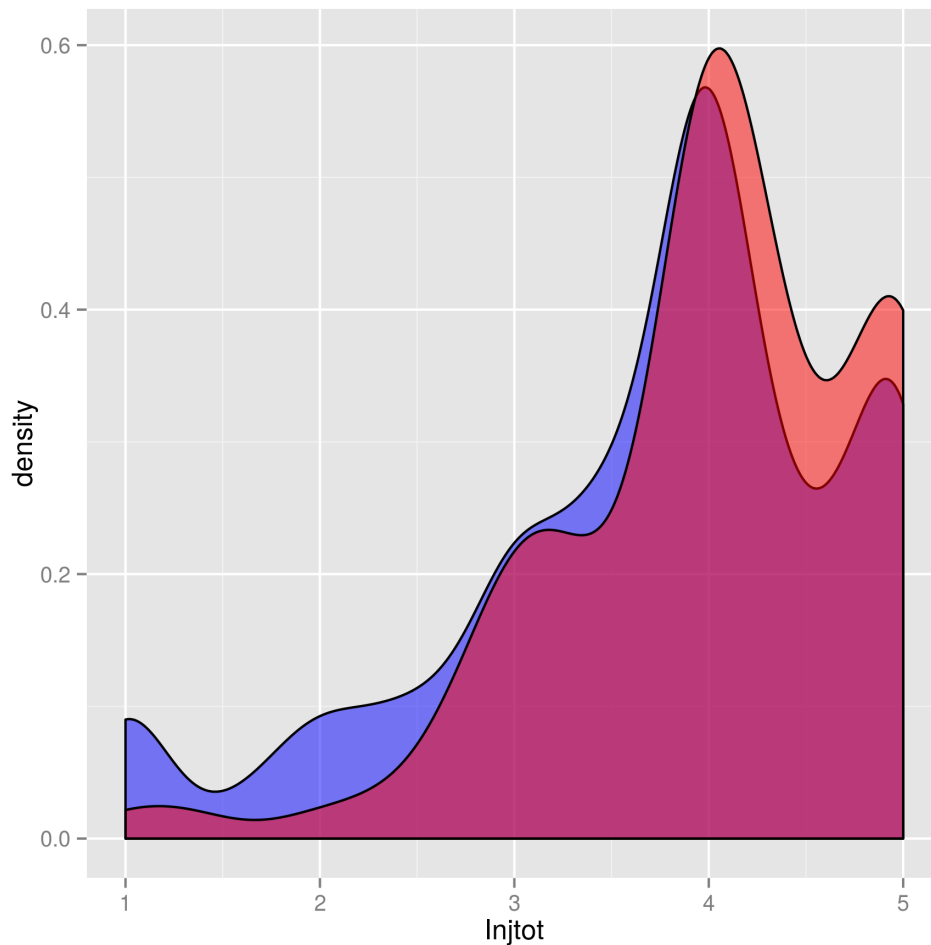


Figure 7.3: Injection Credibility Density Plot, Survey Sample (sample two) (blue), Experimental Sample (red)

analgesic.

Next, the credibility scores for injection painkilling treatments were examined between the survey and experimental groups.

Figure 7.3 shows that the Injection credibility scores were almost identical in their distributions between the two samples.

Finally, the Alternative treatment scores were examined between the two samples. Acupuncture levels in the experimental sample were a little lower than those in the survey sample. In contrast, the credibility scores for Homeopathy were slightly higher in the experimental sample than in the survey sample. While Reiki credibility totals were quite low in both samples, they were a little lower in the survey sample.

7.3.6 Implicit-Explicit Relationships

Next, the relationships between the explicit and implicit measures were examined.

It can be seen from Figure 7.4 that the LOTR was only really correlated with the Acupuncture items and with the MAAS, the Conventional Treatment scales correlated within themselves, as did the Alternative treatment scales, while the two IAT measures showed no appreciable correlations with each other. The relationships between the IATs and explicit measures were small, and surprisingly in the unpredicted direction (negative) . Another surprise was that the direction of the correlation between the LOT-R and the MAAS was opposite to that observed in prior research. Possible reasons for these results are considered in the Discussion, and some theories were discussed in Chapter 4.

7.3.7 Primary Analyses

Using a chi-square test ($\chi^2(1) = 0.32, p = 0.56$) there was no significant difference between the Deceptive and Open Placebo conditions on the probability of placebo response, which would suggest that the suggestion in a each condition was not successful at inducing a response reliably. This is damaging, but not devastating to the main point of the research.

Another Chi-Square test($\chi^2(11) = 64.0, p \leq 0.0001$) showed that priming appears to have quite a large effect, as the Placebo Response by Condition is extremely significant given the priming manipulation.

This finding is much more clearly conveyed in Figure 7.5, where it can be seen that participants were much more likely to respond to placebo following a priming intervention. Given that priming interventions typically take place outside conscious awareness, this suggests that there is at least some part of the placebo response which is amenable to non-conscious (or implicit) influences.

In Figure 7.4 the relationships between the self report and implicit measures are shown. It can be seen that the credibility scores break into conventional and alternative groups, while the LOT-R and MAAS do not correlate hugely with any of the other measures (but do correlate quite well with each other, as noted above).

There was a difference in the mean cream credibility scores by whether or not a participant responded to placebo, but the difference was not significant ($t = -0.9545, df = 53.766, p = 0.3441$).

A number of findings are apparent from the plot in Figure 7.6. The placebo effect was approximately equivalent to a 15% decrease in pain (read from the graph at the point the no response participants pain reached seven). This is a relatively large effect. In addition, the participants who responded to placebo tended to remain in the experiment for a longer period of time (which is intuitively obvious). Below, formal model testing for the major hypotheses takes place.

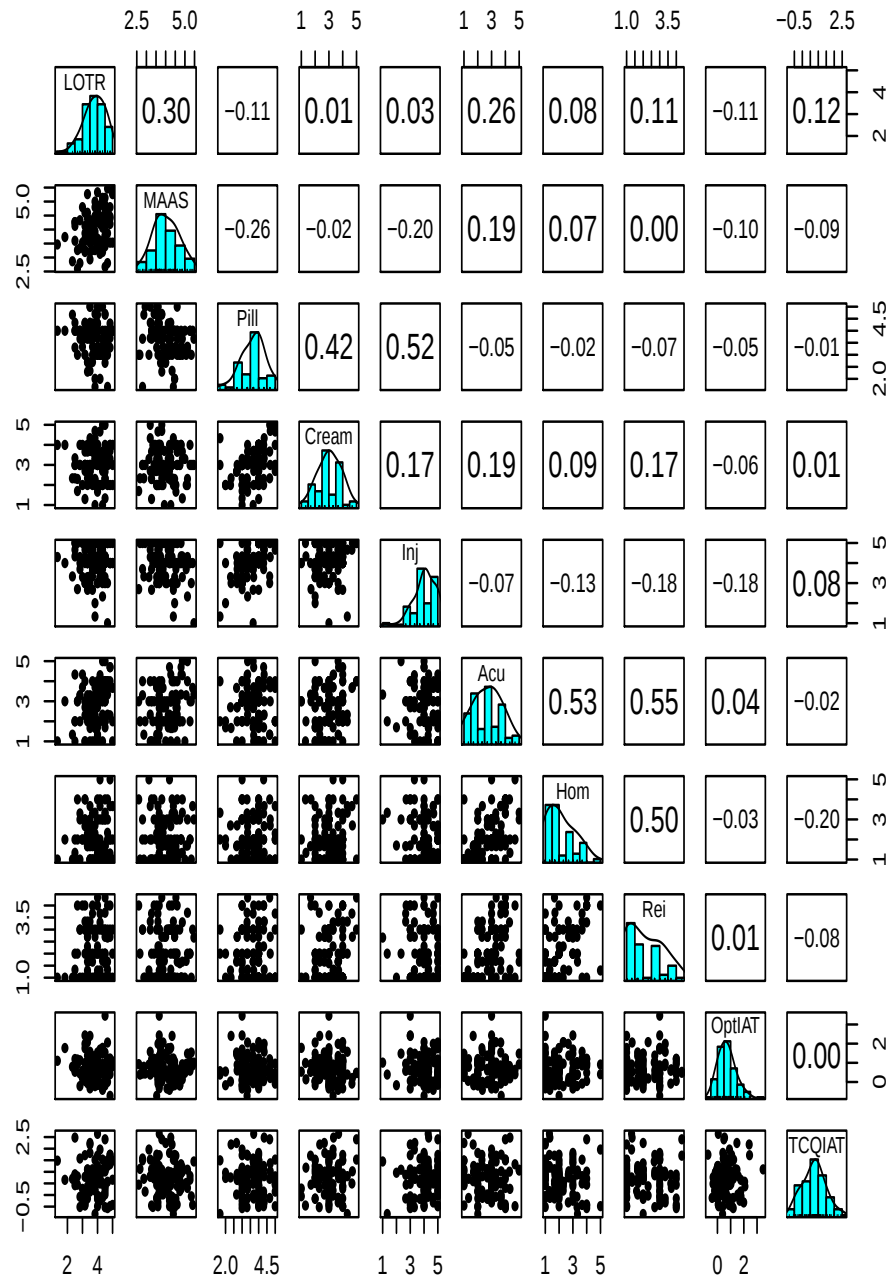


Figure 7.4: Scatterplots and Correlations for Implicit and Explicit Measures. Upper triangle shows the correlations, lower triangle shows the scatterplots, Diagonal shows the distributions. From left top: LOTR - Life Orientation Test, Revised. MAAS- Mindful Attention Absorption Scale. Pill - Pill Credibility. Cream - Cream Credibility. Inj - Injection Credibility. Acu -Acupuncture Credibility. Hom - Homeopathy Credibility. Rei - Reiki Credibility. OptIAT - Mean Scored D Measure, Optimism IAT. TCQIAT - Mean Scored D Measures, Treatment Credibility IAT

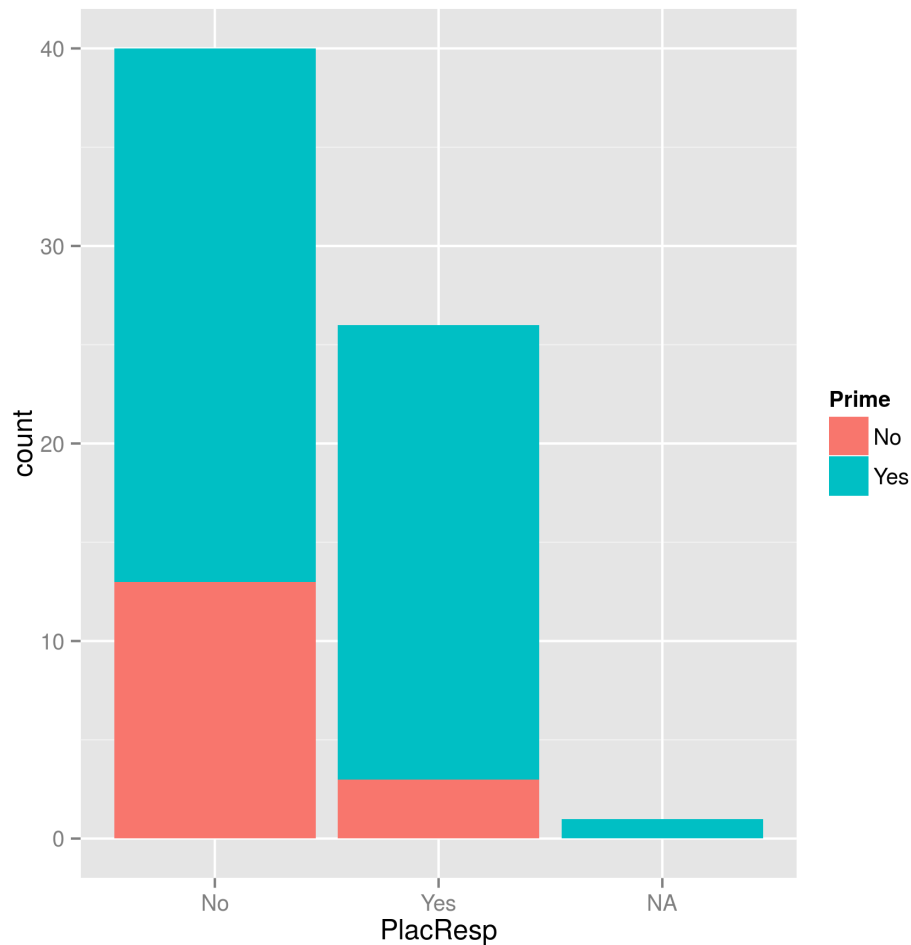


Figure 7.5: Count of Placebo Response in Primed and Non-Primed Conditions

7.3.8 Predicting Placebo Response

In order to examine whether or not the IAT and or explicit measures scores were usefully predictive of placebo response, a logistic regression model was used. The intention was to use a train and test set to ensure that valid inferences could be made after selection of the items, but unlike in previous chapters, this work used a 60% train and 40% test set, to allow more complex models to be tested.

As can be seen from Table 7.6, the model chosen by the stepwise selection procedure has an almost significant coefficient on the TCQ IAT, one for the LOT-R, and the two two-way interactions trended towards significance.

Table 7.7 shows the results of the stepwise selected model on the test data. It can be seen that none of the variables are significant, which is not surprising given the sample size. The odds ratios are a little more informative. These are shown in Table 7.8. As can be seen, there were no significant differences from zero in this sample, suggesting that if there is an effect here, the sample is not large enough to

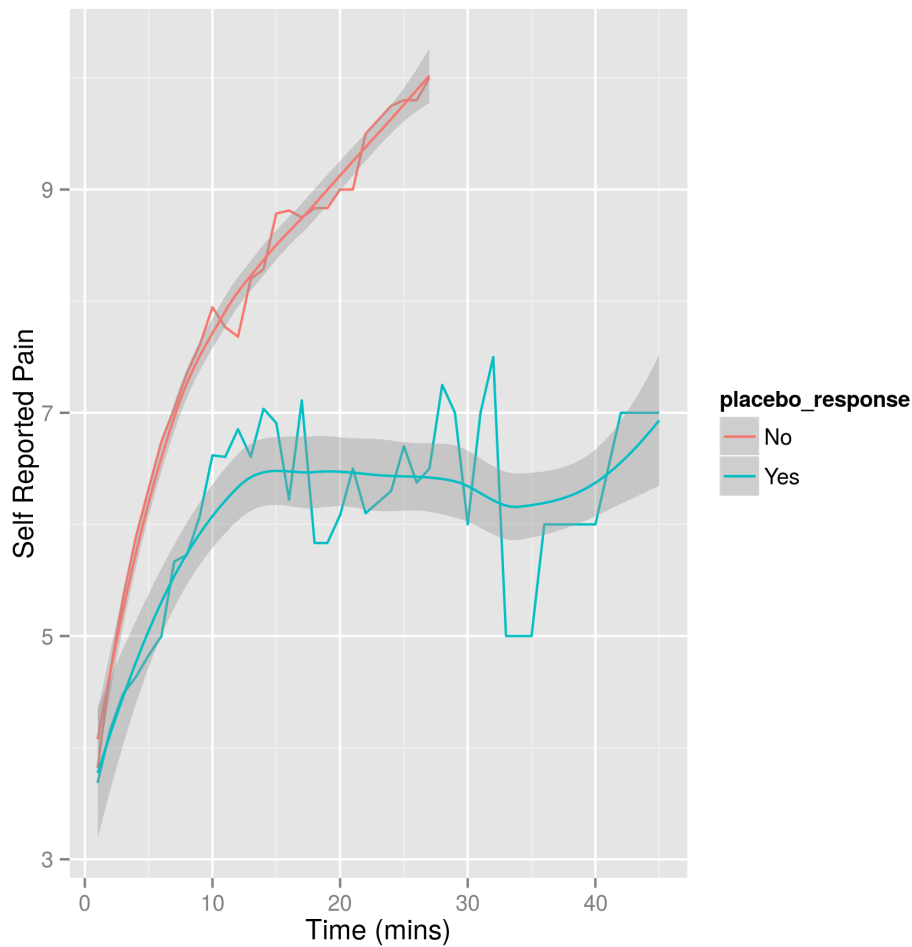


Figure 7.6: Pain Ratings of Participants by Response to Placebo Across Time. Straight line is a Loess smoother, the jagged line represents the actual pain levels

uncover it, and as such the effect must be a small one (as this experiment was powered to uncover only medium or large effects).

To investigate the variability of the estimates, a bootstrapping process was performed over the test data to calculate p-values for each of the coefficients chosen on the training set. The means and standard errors are shown in Table 7.9, and suggest that there is little to no significant relationship between the variables and the placebo response. A bootstrapping procedure over the entire dataset produced results much more supportive of the major hypotheses, but as this analysis does not account for the search process, it is not reported further here.

Individual regressions were run for each of the variables over the training and test sets, but none were significant.

Next, IRT estimated abilities for the LOT-R were tested against the placebo response. Note that the ability estimates were calculated from the reduced five item scale described in Section 4.9.4 on page 96. When these ability estimates are included in

Table 7.6: Model selected by Stepwise Selection over the training data

	B	SE(B)	β	z	Sig.(p)
(Intercept)	-2.53	8.30		-0.31	0.76
TCQIATMean	11.67	7.51	17.70	1.56	0.12
OptIATMean	14.37	8.15	20.74	1.76	0.08
LOTR	0.38	1.94	0.53	0.19	0.85
Acu	0.84	0.56	1.65	1.52	0.13
Hom	-1.12	0.63	-2.21	-1.78	0.07
TCQIATMean:OptIATMean	-22.02	10.06	-33.38	-2.19	0.03
TCQIATMean:LOTR	-2.55	1.78	-3.68	-1.43	0.15
OptIATMean:LOTR	-3.66	2.06	-5.15	-1.78	0.08
TCQIATMean:OptIATMean:LOTR	5.11	2.37	9.96	2.15	0.03

Table 7.7: Coefficients of Stepwise Selected Predictors on Test Data

	B	SE(B)	β	z	Sig.(p)
(Intercept)	-1.47	1.01		-1.45	0.15
TCQIATMean	0.43	0.75	0.86	0.57	0.57
LOTR	0.59	0.70	1.20	0.85	0.40
Age	-3.55	2.28	-7.16	-1.56	0.12
OptIATMean	-0.79	0.91	-1.59	-0.87	0.39
Rei	-0.91	0.65	-1.84	-1.41	0.16
TCQIATMean:LOTR	-1.13	0.69	-2.28	-1.65	0.10

Table 7.8: Odds Ratios for Coefficients in Test Data Model

	Mean	2.5 %	97.5 %
(Intercept)	0.23	-4.28	0.10
TCQIATMean	1.53	-1.01	2.06
LOTR	1.81	-0.73	2.19
Age	0.03	-9.97	-0.34
OptIATMean	0.45	-2.78	0.87
Rei	0.40	-2.35	0.28
TCQIATMean:LOTR	0.32	-2.76	0.09

place of the sum-scores, the result does still not achieve the conventional threshold of significance. These results are shown in Table 7.10.

The next variable to be tested in terms of univariate regressions are the treatment credibility scores. These were tested in both their raw forms and as estimates from the successful IRT models described in Chapter 5.

Table 7.11 shows the estimated coefficients for the logistic regression of all three conventional sum scores on Placebo Response. As can be seen, the results were not significant.

Next, the ability estimates generated through the best fitting IRT model were used as predictor variables in a regression on the placebo response.

Table 7.12 shows the results of fitting the model using the IRT estimated abilities for

Table 7.9: Summary statistics for the bootstrapped p-values

	Mean	SD	Median	Min	Max
(Intercept)	0.67	0.32	0.73	0.00	1.00
TCQIATMean	0.58	0.38	0.58	0.00	1.00
LOTR	0.57	0.37	0.52	0.00	1.00
Age	0.48	0.40	0.26	0.00	1.00
OptIATMean	0.66	0.33	0.72	0.00	1.00
Rei	0.58	0.38	0.57	0.00	1.00
TCQIATMean:LOTR	0.58	0.38	0.58	0.00	1.00

Table 7.10: Regression of Placebo Response on IRT Estimated Optimism Scores

	B	SE(B)	β	z	Sig.(p)
(Intercept)	-0.51	0.44		-1.14	0.25
AbilityLOTR	-0.25	0.35	-0.39	-0.72	0.47

Table 7.11: Regression of Conventional TCQ Summary Scores on Placebo Response

	B	SE(B)	β	z	Sig.(p)
(Intercept)	-2.64	1.79		-1.47	0.14
Pill	0.17	0.52	0.23	0.34	0.74
Cream	0.05	0.33	0.10	0.16	0.87
Inj	0.34	0.38	0.56	0.88	0.38

Table 7.12: Regression of IRT Estimated Ability Scores on Placebo Response

	B	SE(B)	β	z	Sig.(p)
(Intercept)	-0.52	0.27		-1.92	0.06
AbilityConv	0.48	0.41	0.64	1.15	0.25
AbilityAlt	-0.18	0.34	-0.29	-0.54	0.59

the Treatment Credibility Questionnaire. As can be seen, the model is still not significant. However, the AIC was slightly lower for the model fitted using the ability estimates, which suggests that though they are not particularly good predictors, they are at least somewhat better than the raw sum-scores.

Table 7.13: Logistic Regression for the Impact of Mean GSR measurements on placebo response

	B	SE(B)	β	t	Sig.(p)
(Intercept)	-0.63	0.44		-1.44	0.15
Mean GSR	0.02	0.03	0.37	0.71	0.48

As can be seen from Table 7.13, the mean level of skin response was not significantly associated with the response to placebo.

7.3.9 Pain Ratings

As can be seen from Figure 7.7, the Deceptive Placebo group appeared to report lower pain ratings than the Open Placebo group, but from a cursory investigation of

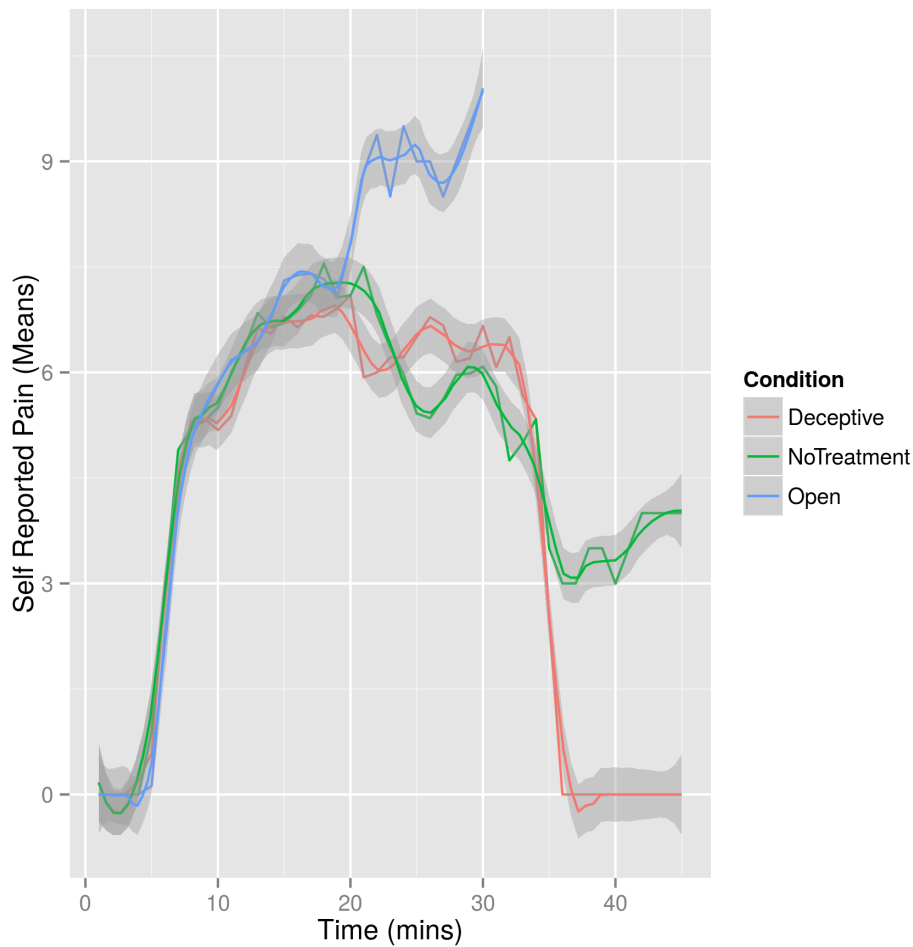


Figure 7.7: Plot of Pain Responses by Condition Over Time

the plot, did not appear to be significantly different to the No Treatment group. The curves shown in the figure used a locally weighted smoother (Loess, span=0.2) to create the lines, given the substantial non linearity of the results. However, this plot does show that there was a significant difference between the placebo group and the two other conditions.

As can be seen from Figure 7.8, the results of the median pain ratings by group show a somewhat different pattern, in that the drop in the pain ratings for the Deceptive Condition is much less apparent, suggesting that the majority of the decrease in pain ratings was driven by a small number of extreme ratings in this group. Again, a locally weighted smoother (Loess, span=0.5) was used to fit the curves.

The length of time participants spent having physiological readings is described in Table 7.14. Note that this time is counted from the start of the experiment, not the application of painful stimulus. This chart explains why the plots of GSR by condition and other grouping variables are not consistent in length.

The mean pain ratings were also examined for autocorrelations and the results were

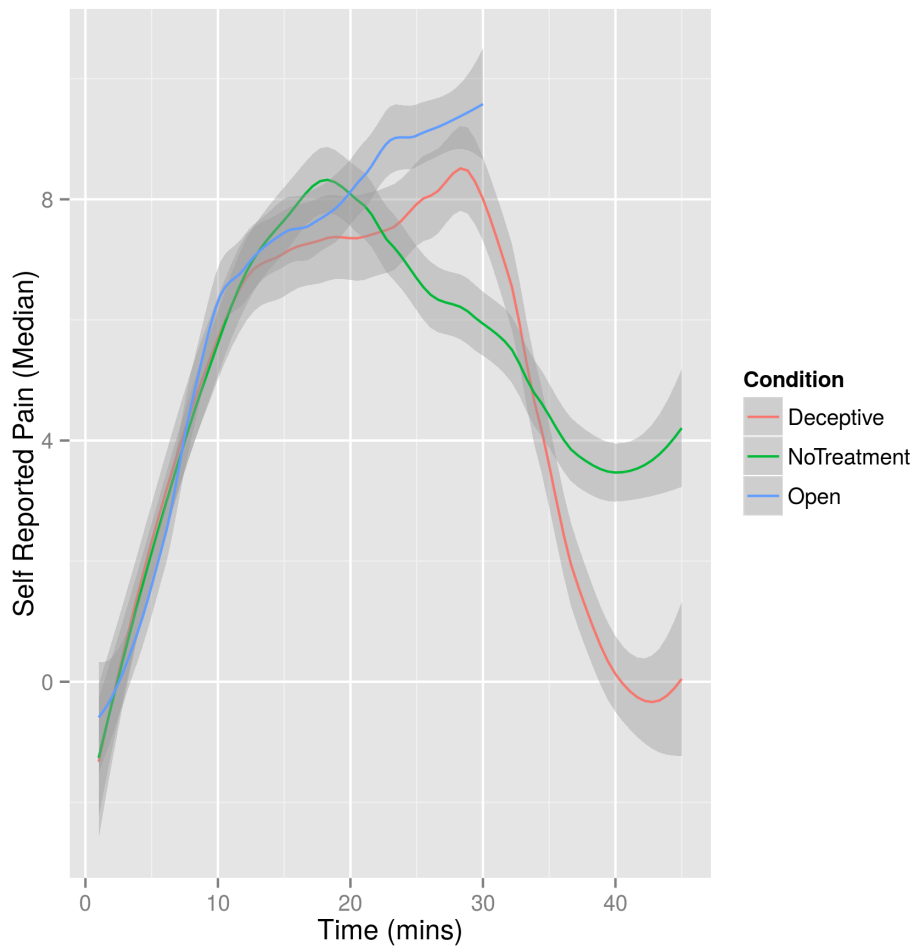


Figure 7.8: Median Pain Ratings Over Time by Condition

Table 7.14: Median Recording Time (secs) by Condition and Placebo Response. Time is counted from the beginning of physiological monitoring (approximately 5 minutes before application of painful stimuli)

Condition	PlacResp	surv	max_surv
No Treatment		1195.51	3187.43
Placebo	No	1169.64	2459.62
Placebo	Yes	1069.42	1430.12
Placebo		641.47	641.47
Treatment	No	1002.91	2350.97
Treatment	Yes	992.58	1887.88

exactly the same as for the median pain ratings, indicating an ARIMA(1,3,1) model was the best fit for the data.

7.3.10 Examining the Effect of Pain on GSR

The next step in the analysis was to examine the impact of the particular painful stimuli administered on skin conductance. Some preliminary research has suggested

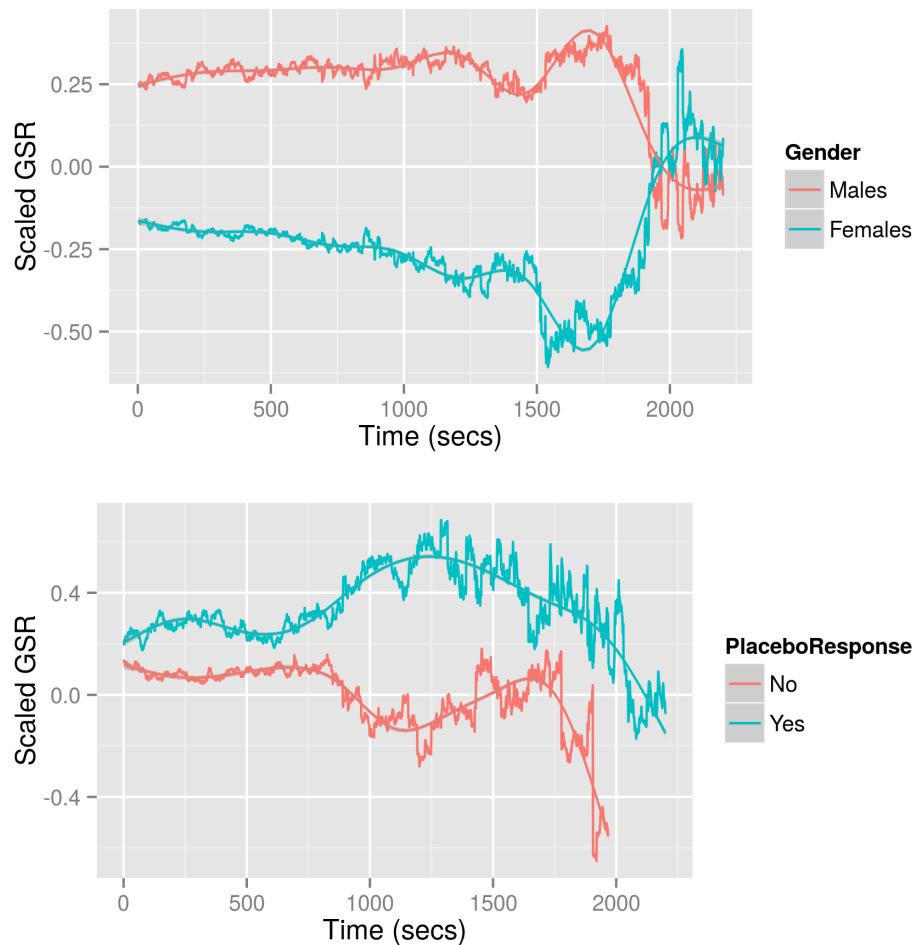


Figure 7.9: GSR over Time by Response to Placebo

that this may be a useful predictor of painful stimuli, and this question will be covered in the following section. However, in this section, the impact of the painful stimuli on skin conductance will be assessed.

Before the analysis of GSR data was conducted, the mean GSR per group were scaled to ensure that they were directly comparable. Scaling was performed using a z-score method, where each observations value was subtracted from the mean and divided by the standard deviation of all the observations.

Figure 7.9 (top panel) shows that the pattern of GSR responses was quite different for males and females. Males showed somewhat decreasing GSR over time (but note that the number of participants in each group fell off as time progressed, so the pattern is less than certain), but an overall higher GSR throughout the entire experiment, while females showed the opposite pattern.

As can be seen from Figure 7.9 (bottom plot), the GSR levels of those participants who responded to placebo increased over time, while the GSR levels of those who did not decreased over time.

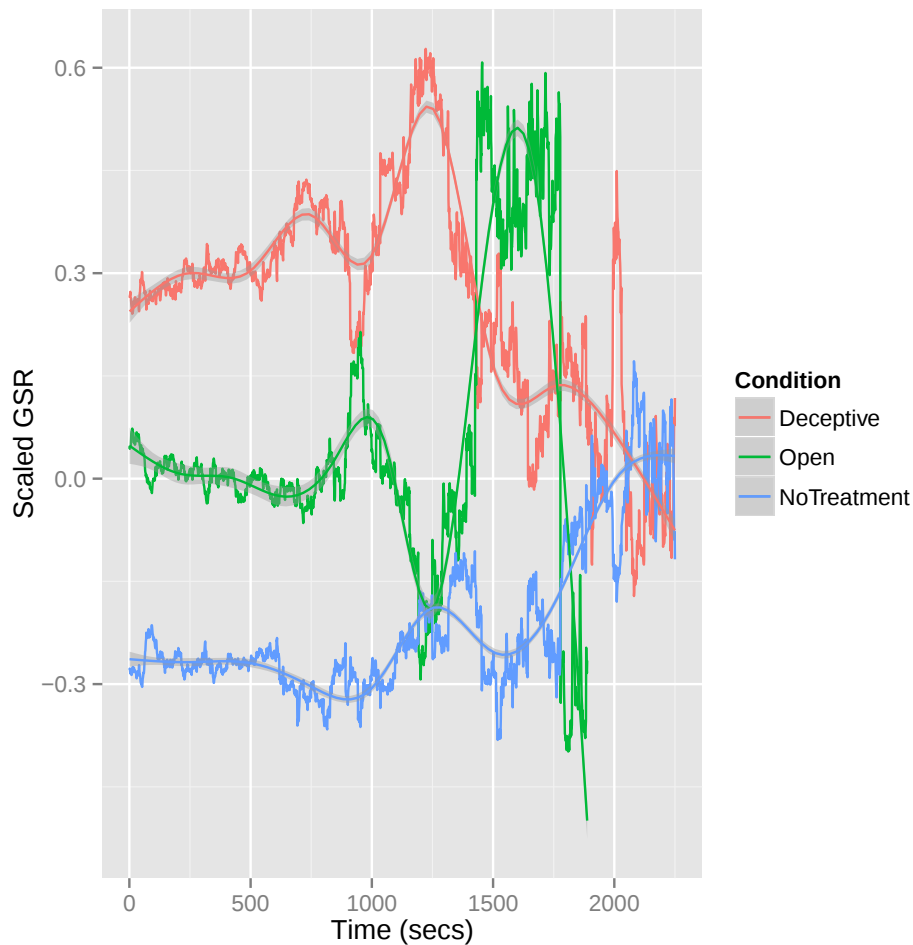


Figure 7.10: GSR by Condition over Time

This seems intuitively odd, as one might have expected the placebo effect to have lowered skin conductance in the group which responded to placebo.

An examination the GSR levels by condition over time (Figure 7.10) showed that the Deceptive Placebo Group show a large increase in skin conductance across the study, while the Open Placebo Group show a parallel decrease in GSR levels. The No Treatment group show a steady increase from low levels. It can be seen that even when the GSR scores are scaled and centred, the groups were not comparable at baseline, which is surprising, given that the groups were balanced on all other measures.

The Open Placebo group had the lowest GSR over time, and that the deceptive pain group seemed to have the highest GSR throughout the experiment. The No Treatment group chart is perhaps the strangest, showing a steady upward trend until about 25 minutes into the experiment, and then falling steadily from there to reach approximately the level which it began¹.

¹Note that these are means by condition, which means that the later values represent less



Figure 7.11: Mean GSR Levels by Condition over Time

This may have occurred due to habituation to the experimental environment, as they had no experimental manipulation during the course of the study which might have changed their GSR. In contrast, the Open Placebo group show a low GSR starting off which then rises slowly, dips and then rises again. The course of GSR in the Deceptive Placebo group looks like what would have been expected from the No Treatment group; i.e. a slow and steady rise throughout the experiment.

For the Deceptive Placebo Condition, the mean cross-correlation was 0.63, for the Open Placebo condition it was -0.53, and for the No Treatment Condition it was -0.22. As can be seen the magnitudes for the two active conditions were similar, though the directions were reversed, while for the no treatment group, the correlations were much smaller, suggesting that there was less relationship between the physiological measures and the pain ratings in this condition. This subject is interesting, and is explored more below and in the Discussion.

A regression which is shown in Table 7.15 showed that there was a significant effect of participants.

Table 7.15: Regression of Condition on Mean GSR

	B	SE(B)	β	t	Sig.(p)
(Intercept)	7.92	1.23		6.46	0.00
ConditionPlacebo	1.66	1.73	0.18	0.96	0.34
ConditionTreatment	3.79	1.82	0.41	2.08	0.04

the Deceptive Placebo Condition on mean GSR scores, suggesting that the mean GSR scores tended to be higher in this condition as compared to the others. This would seem to match the descriptions of the cross-correlation functions above.

7.3.11 Testing the Theoretical Models

7.3.12 Relationships between Explicit and Implicit Measures

Next, three models were fit to examine the relationship between the implicit and explicit measures themselves. Of interest here was whether or not mindfulness (operationalised using the MAAS) would moderate the relationships between explicit and implicit measures.

The three models were as follows:

1. A two factor correlated model between implicit and explicit measures
2. A two factor uncorrelated model with implicit and explicit measures (this model did not converge)
3. A two factor model where the covariances between implicit and explicit measures were moderated by mindfulness.

Unfortunately, none of these models converged, suggesting that either there was not enough data to estimate them, or that they are not useful models.

Table 7.16: Fit Statistics for the SEM Models on the Relationship between Implicit and Explicit Expectancies

	chisq	df	pvalue	cfi	rmsea	rmsea.ci.lower	rmsea.ci.upper
One Factor	10.30	5.00	0.07	0.00	0.10	0.00	0.19
Two Factor	10.10	4.00	0.04	0.00	0.12	0.03	0.22

The next step in the use of SEM models was to examine whether a two factor or one factor model best fitted the relationship between implicit and explicit models.

Comparing a model where both implicit and explicit measures loaded on one factor to a model in which they loaded on separate factors showed that the one factor solution provided a better fit to the data. Given the lack of data for a test sample in this portion of research, these results should be regarded as provisional until replicated.

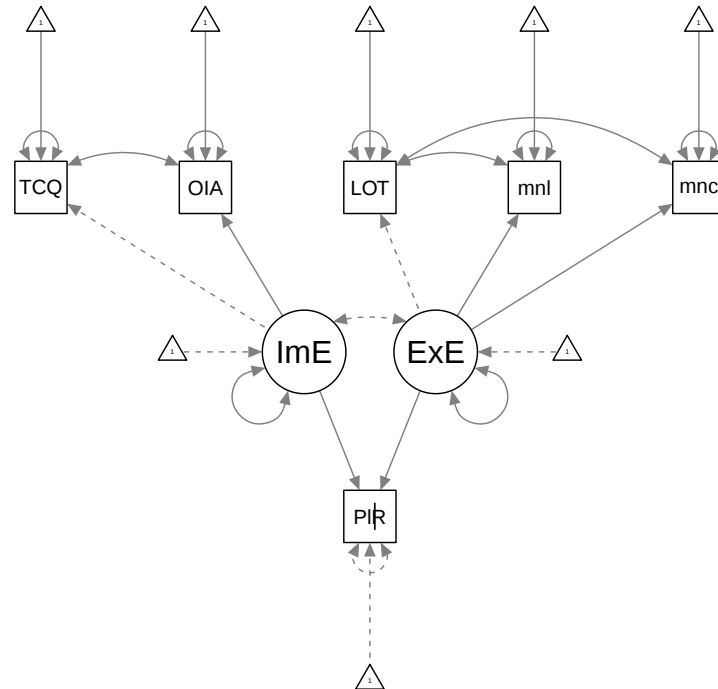


Figure 7.12: Direct Placebo Response Model (Model 1)

7.3.13 Theoretical Models

The final step in the analysis was to test the theoretical models which were proposed for the relationship between the Placebo Response and implicit and explicit expectancies in line with the models described earlier, in Section 7.1.1.

Figure 7.12 shows the direct placebo model, where implicit and explicit expectancy factors have direct effects on the placebo response. This model did not converge, and as such, this model will not be discussed further.

The next model fit is the model of Kirsch, where all other parameters are mediated by (explicit) expectancies and implicit expectancies have no impact.

Next, a model where expectancies and optimism are jointly responsible for the observed placebo response, and implicit expectancies play no role.

Next, a model where optimism was the sole mediator of the placebo response was fitted.

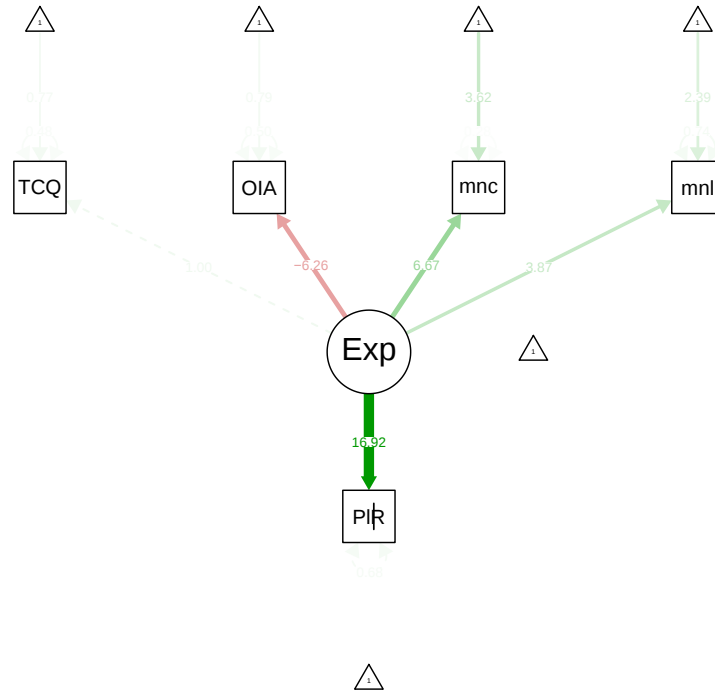


Figure 7.13: Placebo SEM Kirsch with implicit measures

Finally, a model like that of Kirsch, but including the implicit measures was fitted. This model is shown in Figure 7.13.

Another model was fitted with an independent effect of GSR on the placebo response, and all other variables mediated through a general expectancy factor. Unfortunately, this model did not converge, and so is not discussed further.

Table 7.17: Fit Measures for Placebo Response Structural Equation Models

Model	df	pvalue	cfi	rmsea	rmsea.ci.lower	rmsea.ci.upper	gammaHat	adjGammaHat	baseline.rmsea
Optimism	5.00	0.12	0.00	0.11	0.00	0.23	0.96	0.95	0.06
Credibility	4.00	0.60	0.00	0.00	0.16	1.02	1.01	0.06	2.77
Kirsch Two Factor	5.00	0.77	1.00	0.00	0.00	0.12	1.02	1.05	0.00

From examining the fit measures shown in Table 7.17 it would appear that the best fitting model is that of Kirsch when implicit expectancies are included in the model. Implications are elaborated upon in the discussion. Figure 7.13 makes this clearer, as it can be seen that the Optimism IAT (OIA) has a negative weighting on the expectancy variable, while the mean conventional credibility (mcn) and the mean alternative credibility have a positive weighting on the latent expectancy variable,

which impacts the response to placebo. This model would seem to suggest that the explicit factors performed better as measures of expectancy than did the implicit measures. Unfortunately the model with implicit and explicit expectancies did not converge, as the differences in fit between this model and the two factor model would have proved illuminating.

The coefficients for this model are shown in Table 7.18. The t_1 coefficient represents the probability of a positive placebo response. None of the coefficients were significant, in line with the previous regression analyses.

Table 7.18: Coefficients for Kirsch Two Factor Model

	Coefficients
PlacResp~LOTR	-0.16
PlacResp~convaltcomp	0.08
LOTR~OptIATMean	-0.13
convaltcomp~TCQIATMean	0.08
LOTR~~LOTR	0.63
convaltcomp~~convaltcomp	0.94
PlacResp t_1	-0.49
LOTR~1	3.80
convaltcomp~1	1.27

7.4 Discussion

7.4.1 Relationship between Implicit Measures and Placebo

Some evidence did suggest that the IAT measures were related to placebo response. The step-wise selected model was significant on the training data, however this was not replicated on the test data. It is worth noting that the z-values were relatively large here, however, and the lack of significance may be an artefact of sample size. It has been established that given this study, a much larger ($N=150$ in each group) would be needed to have 80% power to detect a difference of this size.

The interaction between the LOT-R and the TCQ IAT is of interest, but given the relatively small sample size here, this could be mere statistical variability.

Secondly, given the number of models fitted, some were almost certain to come up as significant, and the interaction between explicit and implicit measures was not a prior hypothesis of this research.

On the balance of probabilities, based on this research, it can be said that implicit measures are unlikely to prove useful in the prediction of placebo response in healthy volunteers, and so at least one of the theoretical predictions from Chapter 3 has been shown to be incorrect.

7.4.2 Modelling the Placebo Response with SEM

This section threw up a number of interesting results, with the caveat that these results should be regarded as very preliminary. There were some difficulties with fitting these models in general, given the small sample size, and a number of models of theoretical interest did not converge.

However, it does appear that based on the data collected here, that both implicit and explicit measures provide a useful fit to the data. This should be qualified by noting that the lack of significance in the logistic regression models carried over to the SEM analyses.

The results of this study would appear to support the theoretical model of Kirsch, whereby all other variables are mediated by expectancies, which exerts the change upon the placebo response.

However, one difference here is that a model with implicit and explicit measures provided a much better fit to the data than did either of these variables alone, suggesting that the theory of Stewart-Williams is not correct.

However, the best-fitting model used these as inputs to a generalised expectancy factor, which would suggest that the model of Kirsch remains valid. Additionally, the results here would seem to suggest that the variance captured by optimism is a subset of the variance explained by expectancies, at least in this experiment. Given the known situational impact of different constructs on placebo response, this is not surprising, but the relationships between expectancies and optimism had not been investigated in a placebo analgesia study before, so at least this finding has been replicated.

To recap the predictions from earlier, it appears that there is a relationship between placebo response and physiological parameters (see Section 7.4.3), but the exact nature of that response, and its directionality remains to be determined. There was some weak evidence for a relationship between implicit and explicit measures and placebo, but this would require independent replication to be verified. Mindfulness was not related to the placebo response, but did moderate the relationship between explicit and implicit measures.

7.4.3 Physiological Analyses

This section was perhaps the most interesting, in that some observations were made that do not appear to have been reported frequently in the literature, with the exception of some older research of Kirsch (Kirsch & Weixel, 1988).

Firstly, the GSR responses appeared to vary systematically as a result of Condition,

more specifically as a result of the suggestions given for the Deceptive and Open Placebo conditions. This can be seen by examining the cross-correlations between pain ratings and GSR responses. These were systematically correlated in the two “active” conditions, where suggestions were given, while not correlated at all in the No Treatment condition.

Given that the only difference in the experimental conditions was the suggestions, this would seem to suggest that the suggestion can affect the physiological responses of the participants. More interesting was that the two suggestions appeared to have directionally different influences, in that GSR appeared to go down in the Open Placebo condition, while it increased in the Deceptive Placebo condition. However, this finding must be qualified by the finding that GSR levels were different between the groups at baseline and so it is possible that the results here are affected by participant variability.

Finally, it appeared that participants who responded to placebo had a qualitatively different GSR responses to those who did not. Again, this is something which has not been reported in the literature before, and as such deserves further investigation.

Linking this back to the theoretical perspective advanced in Chapter 3, it would seem to support part of the theory, in that the physiological responses differed across conditions. However, the lack of a lagged effect would suggest that either there is no feedback loop between physiological sensations or expectancies, or that the data collected in this experiment was not precise enough to estimate this correctly.

7.4.4 Priming and Placebo Response

The final contribution of this research is that it provides further evidence that priming manipulations can be used to induce placebo effects in a pain paradigm. This is important for the future development of the field, as priming is much less resource-intensive than current procedures of conditioning. This will assist future research as it will allow for more powerful studies to be performed with fewer participants. This may also prove useful as an adjunct to standard clinical practice, subject to ethical concerns being mitigated.

7.4.5 Scoring Methods

A secondary aim of this study was to examine if IRT scoring methods might prove better when compared to sum-scores in the experiment. This prediction was not borne out by the data, although there were some indicators that it might have some merit. Nonetheless, the methodology applied in this research might prove more useful in other research domains.

The relationship between LOT-R and MAAS scores was also of some interest, as it was in the opposite direction from that observed in the previous studies. The notion that the previous results were an artefact of the online scoring method does not hold up to scrutiny, given that the same results were observed in the paper-based scoring method from Study 1 (see Chapter 4). However, one difference between the prior use of these measures with that in the current study was the order in which they were administered. In the previous work, the MAAS questions were answered before the LOT-R, while in this chapter, this order was reversed (as an artefact of the printing process). Given that the MAAS represents mindfulness as a lack of mindlessness, it seems somewhat plausible that answering these questions made the lack of their mindfulness salient to those participants who would typically have rated their optimism relatively highly, and thus moved their scores downwards, accounting for the observed results. This theory is discussed more in Chapter 8.

7.4.6 Conclusions

To recap, the following has been learned from this research.

1. There appears to be a relationship between priming and the placebo response
2. There appears to be no direct relationship between the implicit measures used in this research and the placebo response
3. There may be an interaction effect between the placebo response and a combination of explicit and implicit measures
4. There appears to be a noticeable effect of placebo administration/suggestion on electrodermal activity.
5. A model involving both implicit and explicit measures of optimism and expectancies mediated by a general expectancy factor appears to provide a better fit to the data than do alternative models.

It really needs to be clarified that given the amount of leeway taken with the modeling process in this research that these findings must be regarded as tentative until borne out by future research.

The contributions of this research are as follows. Firstly, a relationship between priming and the placebo response was demonstrated in a placebo analgesia paradigm. This is important for both experimental and clinical reasons. From an experimental perspective, priming has the advantage that it does not require special equipment or a two or three administration design. This will allow for larger placebo responses to be demonstrated under experimental conditions, which will help the field to progress in its aims. From a clinical perspective, this finding also seems to suggest that simple priming of patients by word games (or potentially cues in their environment) could

aid in the healing process. Such a practice would need to be squared with the need to achieve informed consent, but given that the risk of harm from priming seems quite low, this might be acceptable to the general medical community.

Next, the relationship between implicit and explicit expectancies was assessed. No real significant or robust relationship was found here, but given that the effect sizes were rather small, this does not entirely rule out the possibility of an impact, but it does suggest that any effect would not explain large portions of the variance.

Additionally, there was an interesting relationship noted between skin conductance and experimental condition, as well as gender. While the groups were comparable at baseline, both the Deceptive and Open Placebo groups showed significant changes in GSR after administration of the treatment. Given that the treatment was the same in both cases, this would appear to suggest that skin conductance was affected by the suggestions given as part of the experiment. The fact that these changes were in opposite directions may indicate that either GSR was tracking pain responses, or the effect of certain and uncertain expectancies. This would seem to be a prime area for future research on the placebo.

Chapter 8

General Discussion and Conclusions

8.1 General Discussion

The central aim of this thesis was to examine the measurement of treatment expectancies in the context of the placebo effect. This thesis used explicit (self-report), implicit and physiological variables to deliver greater understanding of these psychological constructs and their relationship with the placebo effect.

The thesis was structured into a number of parts, each of which will be discussed in turn, setting the discussions at the end of each chapter into a fuller context. Following this, the overall aims and objectives of the thesis will be reviewed and set into context. Next, the specific contributions of this research to the literature will be elucidated. Finally, some avenues for future research will be described.

The first section of the thesis was the review of the literature around the placebo effect and implicit measures, along with any known individual-level predictors of the effect. This review found that placebo effects have been seen to be mediated by optimism (as operationalised by the Life Orientation Test, Revised (LOT-R) (Scheier et al., 1994)) (Geers, Helfer, et al., 2005; Morton et al., 2009) in both pain and non-pain paradigms. This review also linked the constructs of optimism, which can be defined as “generalised outcome expectancies around the future” (Carver et al., 2010), with the construct of response expectancies of Kirsch (Kirsch, 1985, 1997), which are described as the “expectation of a non-volitional response”. These constructs both crucially rely on the participants’ perception of the likelihood of the event, and a secondary aim of the thesis developed which was to examine the usefulness of these as two separate rather than one combined predictor. The hypothesis was that these two constructs should be significantly correlated, and that one of them would mediate the impact of the other on the observed placebo response (Geers, Helfer, et al., 2005).

Another important potential moderator variable came to light in the literature review, which was the construct of mindfulness, as operationalised by the Mindful Attention Absorption Scale (MAAS) (Brown & Ryan, 2003). This variable was found to moderate the correlations between explicit and implicit measures in an experience sampling study. Furthermore, implicit measures were found to be better predictors of spontaneous responses, while explicit measures were found to be better predictors of deliberate behaviour (Levesque & Brown, 2007), a pattern known as the double-dissociation effect and observed in IAT studies in a variety of domains (Asendorpf et al., 2002; Perugini, 2005; Grumm & von Collani, 2007; Steffens & König, 2006).

This research literature suggested that a good strategy would be to replicate the LOT-R and MAAS on a number of samples from which the eventual experimental sample would be drawn, both to develop better models for these constructs’ relationship (as they had not been concurrently examined when this work was carried

out) and to tailor the measures towards the experimental sample, a strategy which was hoped to increase the precision of estimation on the overall sample. In this research, the RAND-MOS was used both to replicate the impact of these primary constructs on health, and to provide a weak link to health and treatment expectancies in preparation for the experimental research.

One large problem with the development of an implicit measure of treatment expectancies was that there exists no gold-standard measure of explicit treatment expectancies, despite them having been shown to materially affect the outcome of clinical trials (Linde et al., 2007; Bausell et al., 2005; Benedetti, 2005). This was an issue both because typical methods of IAT development rely on such a self-report measure, and because without such a measure, the incremental improvement (if any) granted by an implicit measure would not be clear. Therefore, the development and testing of such a measure formed part of the work of this thesis.

Another fact which became extremely clear during the course of the literature review, is that current methods of both developing and validating implicit measures (or more specifically, IATs), rely extremely heavily on the existence of prior self-report instruments. For the placebo effect which was the focus of primary investigation, no such self-report measure existed, which meant that a new approach needed to be tested. A review of methods in this space revealed the work of George Kelly on personal construct theory (G. A. Kelly, 1991), and more specifically the repertory grid introduced in Book 1 of the aforementioned reference. It was believed that such an approach might prove useful in the current case, and might also prove useful with other constructs of interest in the future¹.

Therefore, in order to develop these stimuli, it was decided to do some qualitative research around health constructs with doctors, alternative therapists and students, to gain an appreciation for how they contextualised and understood issues around health, medicine and sickness. The results of this procedure, along with a parallel exercise asking participants to rate the five most important people or qualities of people related to health, would then be used to develop the Health Repertory Grid.

Finally, the development of the experiment was planned, and it was decided to use a placebo analgesia design, inducing pain with the sub-maximal tourniquet technique (Moore et al., 1979), as this form of pain induction has been shown in a meta-analysis to produce the largest effect sizes (Sauro & Greenberg, 2005).

Below, the results and findings from each of these sections are described and placed in a context of larger research.

¹I am indebted to Drs Jurek Kirakowski and Sean Hammond for originally suggesting this idea.

8.2 Health, Optimism and Mindfulness

This study consisted of approximately fifteen hundred ($N \approx 1500$) observations, of which 1100 were collected in an online fashion, and four hundred were collected in pen and paper format.

The major aims of this part of the research was to collect useful background data for the experimental sample, enabling psychometric models to be built on large samples, which could then be applied to the experimental sample, as well as developing psychometric models which could tease out the relationships between the health-related constructs (RAND MOS), optimism (LOT-R) and mindfulness (MAAS). The headline finding of this research, which was replicated across both samples (and a further one which was not analysed for this thesis) was that optimism was negatively related to both health and mindfulness, a finding which has not been reported before.

This finding can be better understood in the light of some of the SEM results on the same dataset. In these, it was determined that emotional well-being moderated the impact of this surprising effect.

Additionally, this strange phenomenon replicated across both original studies, and also in the pilot experiment. However, it did not replicate in the main experimental study, raising the issue of what the critical differences between the samples were. Upon examination of the materials, the only difference between the setup for the experimental study was the order in which the measures were administered. In all the previous studies, the MAAS was filled out before the LOT-R, while in the experimental study, the order was opposite. There are not many examples of these kinds of effects in the psychological literature, but some survey research suggests that the order of administration of questions within a survey can have similar impacts (N. Schwarz, 1999).

This finding links in with the other finding that emotional well-being (a sub-scale of the RAND-MOS) moderated the effect of optimism on health. To see this, merely look at the questions which make up the Emotional Well-Being scale, and note its similarity to the LOT-R. It appears that what occurred is that participants who would typically respond highly to optimism (as evidenced by their Emotional Well-Being score) had their response patterns altered by the act of completing the MAAS. This may have occurred because the MAAS operationalises mindfulness in terms of negative mindfulness (i.e. mindlessness). This may have caused the participants high in Optimism to have realised that some of their optimism was unjustified, which therefore depressed their responses on the LOT-R, leading to the unexpected negative correlation with health. However, this is somewhat circumstantial given that this hypothesis was not directly tested in this research, but it was supported by the SEM

carried out in Chapter 4 which pointed towards mindfulness and emotional well being as mediating the strange optimism-health relationship.

Additionally, the cross-validation approaches employed in this section of the research showed an interesting pattern. The factor models which were developed using the reduced scales from IRT analyses showed better predictive ability on unseen data than did those which were developed from factor analysis alone. This suggests that an interesting approach to the development of scales would be to use IRT (specifically Mokken analyses) to reduce the number of items, and use this reduced scale in SEM to determine if this effect is widespread or merely circumstantial to the samples collected here.

8.3 Explicit Expectancy Measure

The next step in the research was the development of an explicit expectancy measure, which would capture the multi-dimensional nature of treatment expectancies more fully (Stone et al., 2005). Again, this measure was developed across two samples, a small first sample which aimed to validate the measure in terms of reliability, and a second larger study which incorporated the results of the first study and utilised the Beliefs About Medicine Questionnaire (BMQ) (Horne et al., 1999) to provide evidence of convergent validity for the measure.

The results of the first study showed that the measure had extremely good split-half reliability ($\alpha = 0.9$). Another finding was that the first instrument broke down in factor analytic terms in terms of Conventional (Pills, Creams, Injections) and Alternative (Acupuncture) treatments. This was obvious both from the intercorrelations of each of the questions, and from the results of factor analytic and IRT modelling. This suggested that two other alternative treatments should be added to the questionnaire, to determine if this was a valid structure and to provide balance.

The second study showed the expected correlations with the BMQ measure (positive for alternative treatments, negative for conventional treatments) which suggested that the measure did have convergent validity. Additionally, the notion of conventional and alternative treatments being separate factors was replicated across both samples, and using factor analytic and IRT methods. Again, the back-testing strategy employed in the previous chapter, using a combination of IRT and FA approaches, performed better than either of these approaches alone, suggesting that this may be a useful methodology in general.

8.4 Repertory Grids and Experimental Pilot

The final piece of research conducted before the main experimental portion was the development of the repertory grid and the development of the IAT measures, along with the piloting of the experimental procedure.

From the coding and analysis of the interviews (see Appendix B), it was determined that the major contrast between styles of health-care (conventional and alternative) was that the alternative therapists regarded themselves as treating a person, whereas the GP's focused more on the symptoms which this person presented with. Additionally, the GP's seemed to focus more on external causes (pathogens, environment) whereas the alternative therapist focused more on internal causes (body-mind linkages, thought patterns). Note that both of the groups focused on the nature of health as both a social and personal responsibility, but that they disagreed on where the responsibility for changes lay. The students in the sample provided either a moderately alternative viewpoint, or a moderately conventional viewpoint, but in no case were they as extreme as any of the practitioners. The major differential theme appeared to be that of conventional versus alternative, and the truth or falsity of health claims, and this was incorporated into the development of the Health Repertory Grid (and later into the design of the Treatment Credibility IAT).

The next step was the development of the Health Repertory Grid. This was developed analogously to the original repertory grid, using health related constructs taken both from the interviews and from the questionnaire asking people to name the five most important health-related people or qualities in their life. Unfortunately, it became apparent in the course of piloting this instrument that participants did not have enough examples of these constructs to enable them to complete this task. Therefore, the treatment credibility IAT was developed using the notion of true versus false and conventional treatments versus alternative treatments, while the optimism IAT was developed from the LOT-R, in line with current practice in the field.

Next, the two implicit measures were administered to a small sample of volunteers, and showed the expected correlations (small, but in the right direction) with the explicit measures of the same constructs.

Additionally, when the experiment was piloted, there was a small effect of the IAT's on the probability of placebo response in a logistic regression. This pilot study also showed that the two IAT's correlated quite well with the respective explicit measures, and that a placebo response could be induced with the procedure. However, given that a majority of the participants did not respond to placebo, it was decided to use a priming procedure to increase the probability of placebo response (Geers, Weiland, et al., 2005).

8.5 Experimental Research

The final part of the thesis was the experimental research. One hundred and eleven (N=111) participants were assigned to either a Deceptive (told drug, got placebo), Open (told placebo, got placebo) or No Treatment Condition. Additionally, participants were randomised to either a Prime or No Prime condition. condition showed a greater placebo response than those who were in the suggestions focused on “clinically proven” (in the Open condition) and “recently approved” in the Deceptive condition, this may have had an impact. There was a significant Prime by Condition interaction, suggesting that priming was differentially impacted by by the condition (Open Placebo participants showed a greater propensity to respond to suggestion after being primed). This would seem to have potentially large implications for medical practice, if replicated.

With respect to the major hypothesis of the thesis, no direct impact of the explicit or implicit measures was shown to be significant in a logistic regression, using a step-wise approach with a training and test set to ensure that the estimated p-values were unbiased. Note that essentially, while this study provides some preliminary evidence for an impact of the TCQ IAT on placebo response, this effect was not significant on the test set (though the z-value was quite large), and so further research would need to be a larger study to confirm or refute this effect.

Finally, the relationship between skin conductance and pain ratings was examined. This analysis showed that for the two treatment conditions, entirely different GSR responses were shown. This would seem to indicate that the difference between certain and uncertain expectancies can be examined using simple GSR equipment as opposed to expensive fMRI equipment. There is very little earlier work on this in the literature, with the exception of one study (Fujita et al., 2000). However, this study showed that the GSR measures changed when an active drug was administered, and did not when a placebo was administered, suggesting that these effects may be different. This finding is a contribution to the research in that it is novel, and if replicated could provide an interesting back-up of self-reported changes, and may help to elucidate the relationship between certainty of expectations and placebo response.

In conclusion, there was no significant relationship shown between implicit measures and placebo in standard analyses, suggesting that any effect, if it exists, is likely to be quite small.

8.6 Strengths and Limitations

In order for this work to prove useful to future researchers, the strengths and weaknesses of the overall approach need to be outlined. Firstly, this research benefited

from the following:

Background data was collected for all self-report measures from the sample population as the experimental sample was drawn. This is extremely useful for two reasons. Firstly, it allowed for the construction of psychometric models from the large samples which could then be applied to the smaller experimental samples. This approach was hoped to increase the predictive power of the self-report instruments used. While this approach did not pay dividends in this research, conceptually it seems like a useful addition to mixed survey and experimental design within psychological research.

Secondly, this background data allowed for an assessment of how the experimental sample differed from the survey sample(s). This is extremely important given that the assumption underlying generalisation of the study is that the sample is random. Disregarding the impossibility of random selection from a self-selected audience (in all cases), it is still useful to examine how different samples from the same population collected for different purposes are similar and different from one another. These techniques are not commonly used in research, and as they proved useful in this research, it is the authors hope that they might prove useful to others.

Another strength of this research (further outlined in Section 8.7.3) is the pervasive use of cross-validation and out of sample testing of all psychometric and regression models. This procedure improves test accuracy by an order of magnitude in some settings (Friedman et al., 2009), and is best at preventing errors when models grow complex, which many psychometric models do. Additionally, this process allowed for efficient use to be made of both survey and experimental data, which is something which makes this approach a useful one for future researchers to take.

In terms of the overall approach taken to the experimental design, this study is one of the first to examine self-report, implicit and physiological variables and their relationship to the placebo response. This is a novel approach for the study of placebo, and one which is important given that the placebo effect straddles the boundaries of the physical and psychological. There were intriguing correlations between the condition and the physiological responses measured, which have not previously been reported in the literature (with the exception of some tangential research carried out by Fujita (Fujita et al., 2000)). This is a fruitful area of study, especially in the light of research suggesting that dopamine levels reported from neurological studies appear to co-vary with the certainty of reward. The only difference observed between the two conditions in this study was the level of uncertainty, and so it is justified as regarding this (until further evidence is acquired) as evidence of the same phenomenon.

In terms of limitations to this study, there are quite a few. The first, and most glaring is that all the research was carried out on students and staff from one university, and it is difficult to know how these results will generalise to other populations given the WEIRDness of the particular sample (Henrich, Heine, & Norenzayan, 2010).

Additionally, another limitation of this research is that the IAT measures were not administered to a large, representative sample. This limitation was due to the unavailability of a computer-assisted methodology to administer these studies across the internet.

Additionally, a further limitation with respect to the development of the IAT's is that the rep grid approach was not successful, which made it impossible to pursue the principled approach to IAT development proposed at the outset.

A final limitation of this research was that it was only powered to detect medium sized effects between groups (in the schema of Cohen (Cohen, 1988)). Given the lack of significance of the main hypotheses in this sample, it may be that implicit measures have no predictive power, or that this research was not powerful enough to detect them. This is a determination which needs to be made by future research in the field.

8.7 Overall Contributions

The overall contributions of this thesis can be divided into three distinct contributions: those which are applicable to clinical research, those which are applicable to experimental research, and those which are methodological. Each of these sets of contributions are outlined below, and in the next section further research within each of these areas will be proposed.

8.7.1 Implications for Clinical Research & Practice

The largest contribution for clinical practice may be the finding that semantic priming can impact the response to placebo. While this has previously been reported in the literature (Geers, Weiland, et al., 2005; Jensen & Karoly, 1991), this study is one of the first to demonstrate that the effects of priming carry over to a situation where the participant is told that they will receive a placebo. This links up with the research of Kirsch & Kaptchuk (T. Kaptchuk et al., 2010) where participants who were given open placebos in an RCT of irritable bowel syndrome showed significant improvements over a no-treatment group which was matched on all interventions except for the placebo. The effect sizes for this intervention were quite small, however, and this research indicates that through the use of priming, these effect sizes can be significantly increased. This would seem to provide a new way for physicians to ethically increase the response of patients to particular treatments, without breaking informed consent.

One difficulty with this approach would be that it could be argued that there is no informed consent for the priming intervention, but if these kinds of semantic priming techniques could be generalised to be a routine part of clinical practice, (perhaps by

being embedded in consultation documents), and the benefits were clear, then these might be seen to outweigh the risks. However, such a decision would be contingent on further research to outline both the benefits and risks of this approach in more naturalistic settings. A proposal for this research is outlined below.

would finding, as trials, and removed using these then the success rate of with actual effectiveness placebo responses in a subset

8.7.2 Implications for Experimental Research

There are also a number of implications for experimental research arising from the work reported in this thesis

The priming finding noted above in the clinical contributions section could also prove extremely useful in inducing more repeatable placebo effects with smaller samples. This is especially germane in the field of pain, given that more participants than necessary for the purposes of the research should not suffer noxious stimuli. Current practice in much of the field is to use a pre-conditioning approach, but this requires more expensive equipment and particular forms of pain induction. Should the priming findings be replicated more widely, this constraint would be removed, and both ethical and replicability concerns within the field could be alleviated.

8.7.3 Methodological Contributions

Another contribution of this thesis was in the general methodological approach taken. As described throughout the thesis, a number of methods were employed to develop the instruments in this thesis which are not commonly used in psychological research.

The first contribution of this part was the collection of survey samples of all the explicit measures used in the experiment from the same population as the experimental participants were sampled from. This is novel in the literature, and provides a useful check on the generalisability of the measures to this population, and insight into how the experimental sample differed from the survey samples.

Another contribution was the pervasive use of cross-validation throughout both the psychometric analyses and the regression modelling. This is a better approach in terms of minimising the number of false decisions made as a result of over-fitting, and its use throughout this work has ensured that all models were tested on unseen data, so it is somewhat more certain that they reflect a real process of interest rather than sample variability.

These two contributions are important in that this thesis can provide a case study for future work which can reuse these tools.

Additionally, the results of the optimism-mindfulness correlations across this thesis suggest that more attention should be paid to the order in which questionnaires are administered. Given that many surveys are delivered across the internet, randomising the order of complete instruments would seem to be a reasonable precaution to ensure that order effects do not lead to spurious correlations (and to assess if any previous studies have been impacted by this effect).

8.8 Further Research

There are a number of useful pieces of further research that would illuminate further some of the research conducted as part of this thesis.

Firstly, the limits of semantic priming need to be assessed in an RCT. It seems logical, given the effects of expectancies (Bausell et al., 2005), that there would be an impact, but the size of this effect would need to be assessed. The difficulty with such an RCT is that given the results of this study, the participants in the placebo group would tend to respond more than those in the treatment group, which might make this study more difficult to run (as no pharmaceutical company would be interested). Nonetheless, this would provide better insight into under what conditions semantic priming has an impact on response to verum treatment versus placebo.

Another piece of further research would be to replicate the GSR-expectancies link described above. If placebo/verum drug conditions can be separated by the use of skin conductance measures, then this would greatly aid in the assessment both of the size and the duration of placebo effects. Additionally, this link may provide important insights into the physiological impacts of placebo.

A final piece of further study which naturally arises from this research is the assessment of order of administration effects across particular self-report instruments. This is extremely important, as it is simply not known to what extent they can affect outcomes. While there has been some research into the norms surrounding questionnaires and the impact of the order of questions, there is much less research on the order effects between self-report instruments, and this research is important to ensure that one measure is not impacting the response on another in unforeseen and unwanted ways.

8.9 Conclusions

This thesis set out to achieve the following:

- To develop and test an implicit measure(s) to predict the placebo response;

- To develop and test a more substantial explicit expectancy measure for use in future placebo-related research;
- To test a number of theoretical models around the relationships between explicit, physiological and implicit measures and the response to placebo;
- To apply better models and methods to the experimental and survey data-sets used.

At this point of the thesis, it remains to assess to what extent these goals have been achieved. Briefly, there was weak to some evidence for a potential effect of implicit measures on the response to placebo. This should be tempered with the fact that a direct significant effect only occurred in a training sample, and did not replicate on unseen data. Additionally, the impact was only significant in the presence of interactions, and as such may represent random sample variability.

Next, the explicit measure developed in Chapter 5 and used in Chapters 6 and 7 appeared to possess both face and construct validity. However, in the experimental portion of the research it did not appear to possess any predictive validity, which would seem to suggest that this measure is not particularly useful in the prediction of placebo response in healthy volunteers.

However, this measure could still prove useful in assessing treatment-related expectancies in general. The instrument has been revised in this thesis to have a common scale for all questions, and additionally has been subjected to rigorous psychometric validation. This should allow the measure to be re-used by researchers interested in the manner in which treatment-related expectancies are conceptualised by participants from the general population. It may be that the measure itself would prove useful if the size of placebo effects had been larger in the experimental study, and this is something that could be tested by future research.

In terms of the testing of theoretical models, a model including both implicit and explicit measures provided a better fit to the data than did either alone. Importantly from a theoretical perspective, a model with a generalised expectancy factor (from Kirsch) provided the best fit to the data. This should be contextualised with the finding that the LOT-R interacted with the TCQ IAT to achieve significance on the training set.

An interesting and unexpected finding from this research was that the GSR of participants appeared to be affected by the Condition in which they were in. In fact, the cross-correlations between the pain ratings and GSR were differentially directional between the Deceptive (positive) and Open (negative) placebo conditions. This would seem to suggest that certain versus uncertain expectancies can impact physiological variables differently. This is an interesting finding, and one which has not been reported before.

So, in conclusion, this thesis has provided a comprehensive test of the predictive ability of implicit measures (IAT's) and a new treatment credibility self-report item. On balance, neither of these measures provided useful predictive power in a large study of placebo analgesia in healthy volunteers.

While the major hypothesis of this study was not demonstrated to be true, some other contributions to the literature have been made

- Priming exerts significant effects on the response to placebo;
- GSR appears to correlate appropriately with response to placebo;
- Models using both IRT and FA approaches perform better than models trained with either method alone;
- The relationship between optimism and mindfulness appears to vary based on the order in which each of the instruments is delivered.
- An approach using cross-validation and replication provides better evidence for the models tested

Appendix A

Appendix 1: Supporting Tables and Figures

A.1 Health, Optimism and Mindfulness Data Appendices

A.1.1 RAND MOS Psychometric Analyses

Table A.1: Factor Loadings, RAND MOS Two Factor Solution, Sample One

PhysEmHealth	PhysicalFunc	Communalites
0.45	0.03	0.21
0.09	0.03	0.01
0.16	0.33	0.16
0.03	0.74	0.56
0.01	0.73	0.53
0.04	0.66	0.45
-0.09	0.91	0.8
-0.01	0.74	0.54
0.01	0.8	0.65
-0.03	0.81	0.64
-0.1	0.91	0.79
-0.09	0.86	0.71
0.25	0.16	0.11
0.3	0.11	0.12
0.19	0.23	0.11
0.33	0.14	0.15
0.58	0	0.33
0.62	-0.1	0.36
0.48	0	0.23
0.73	0.05	0.55
0.34	0.08	0.14
0.36	0.09	0.16
0.64	-0.13	0.39
0.54	-0.06	0.28
0.75	0	0.56
0.69	-0.09	0.45
0.7	-0.1	0.47
0.75	-0.02	0.55
0.6	0.01	0.36
0.74	-0.02	0.54
0.56	-0.07	0.3
0.59	0.07	0.38
0.35	-0.01	0.12
0.42	-0.05	0.17
0.36	-0.04	0.13
-0.53	0.08	0.26

Table A.1 shows the results of the two factor solution for the RAND-MOS from Chapter 4.

Table A.2: Factor Correlations, RAND MOS Two Factor Solution, Sample One

	1	2
1	1.00	0.27
2	0.27	1.00

A.1.2 MAAS Psychometric Analyses

A.1.2.1 MAAS Three Factor Solution

Table A.3: Factor Loadings, Three Factor Solution, MAAS, Sample One

LackAware	PhysUnaware	LackPresAttention	Communalites
0.02	0.22	0.38	0.28
-0.05	0.06	0.6	0.36
0.04	0.11	0.64	0.53
0.13	0.46	0.07	0.35
0.14	0.39	0.07	0.28
-0.03	0.41	0.04	0.17
0	0.84	0.05	0.74
0.35	0.37	0.19	0.6
0.26	0.44	0.01	0.42
0.77	0.17	-0.14	0.68
0.46	0.01	0.16	0.32
0.62	0	0.02	0.39
0.42	0.01	0.25	0.36
0.47	0.12	0.32	0.6
0.52	-0.22	0.24	0.3

Table A.4: Factor Correlations, MAAS Three Factor Solution, Sample One

	1	2	3
1	1.00	0.78	0.67
2	0.78	1.00	0.64
3	0.67	0.64	1.00

Table A.5 shows the five factor solution for the MAAS for Split B.

Table A.6 shows the five factor solution for Split C.

Table A.7 shows the two parameter GRM for the MAAS scale, Sample One.

Table A.8 shows the estimated parameters for the MAAS two parameter GRM for Split B.

Table A.9 shows the estimated parameters for the MAAS two parameter GRM for Split C.

Table A.5: Factor Loadings, MAAS Five Factor Solution, Sample Two, Split B

Distractability	LackPresAware	LackSomaticAware	Inattention	LackPresFocus	Communalites
0.16	0.39	0.13	0.06	-0.04	0.37
0	0.7	-0.01	-0.01	0.08	0.51
0.21	0.48	0.14	0.25	-0.18	0.68
-0.04	0.02	0.83	0.03	0	0.69
0.38	0	0.38	-0.08	0.02	0.39
0.39	-0.18	-0.15	0.21	0.03	0.17
0.89	0.03	0.01	0.02	-0.05	0.82
0.67	0.16	-0.02	0.02	0.16	0.73
0.39	0.04	0.24	0.11	0.11	0.5
0.4	-0.04	0.26	0.04	0.33	0.66
0.18	-0.08	0.17	0.25	0.35	0.48
0.13	0.27	0.05	0.09	0.3	0.4
-0.02	0.01	0.01	0.84	0	0.7
0.16	0.18	0.07	0.35	0.33	0.68
-0.02	0.25	0.05	0.1	0.39	0.35

Table A.6: Factor Loadings, MAAS Five Factor Solution, Sample Two, Split C

LackPresFocus	LackPresAware	LackAtt	LackSomAware	LackAware	Communalites
0.04	0.12	0.18	0.46	-0.17	0.39
0.06	-0.07	0.52	0.19	-0.02	0.39
0.02	0.29	0.59	0.09	0.11	0.74
0.05	0.09	0.01	0.33	0.34	0.4
-0.01	-0.03	-0.02	0.7	0.05	0.48
0.01	-0.2	0.18	0.14	0.39	0.25
0.35	0.12	0.08	0.13	0.42	0.67
0.45	0.09	0.12	0.09	0.32	0.63
0.44	0.35	-0.16	0.16	-0.01	0.51
0.84	-0.01	-0.09	0.06	0.06	0.73
0.44	-0.01	0.16	0.05	0.05	0.33
0.58	-0.07	0.24	0.03	-0.16	0.44
-0.02	0.78	0.06	0.03	0	0.65
0.49	0.24	0.3	-0.05	0.08	0.69
0.26	0.12	0.2	0.12	-0.27	0.26

A.1.3 LOTR Psychometric Analyses

Table A.10 shows the loadings for the LOT-R Two factor solution, Split B.

Table A.11 shows the estimated loadings for the LOT-R Two Factor solution, Split C.

Table A.12 shows the estimated coefficients for the Two Parameter GRM on the LOT-R for Sample Two, Split B.

Table A.13 shows the estimated coefficients for the Two Parameter GRM on the LOT-R, Sample Two, Split C.

A.2 TCQ Thesis

A.2.1 TCQ Factor Solutions

Table A.14 shows the average of the five factor solutions across Splits B and C.

Table A.7: Coefficient Estimates for MAAS Two Parameter Graded Response Model

$\beta^1(se)$	$\beta^2(se)$	$\beta^3(se)$	$\beta^4(se)$	$\beta^5(se)$	α
-3.55	-2.41	-0.97	-0.23	1.11	0.99
-2.74	-1.72	-0.45	0.33	1.68	1.32
-1.90	-0.70	0.20	0.84	2.08	1.30
-2.99	-1.54	-0.65	0.05	1.45	1.17
-2.17	-1.22	-0.43	0.29	1.27	2.24
-2.35	-1.31	-0.47	0.25	1.42	2.50
-2.47	-1.49	-0.61	0.11	1.33	1.67
-2.28	-1.34	-0.40	0.28	1.38	2.23
-2.81	-1.36	-0.01	0.90	2.63	1.18
-3.23	-1.98	-1.04	-0.55	0.34	1.31
-1.74	-0.64	0.35	1.01	2.06	1.36
-2.12	-1.17	-0.23	0.43	1.56	2.22
-2.98	-1.75	-1.03	-0.41	0.66	1.01

Table A.8: Coefficient Estimates for MAAS Two Parameter Graded Response Model, Split B

$\beta^1(se)$	$\beta^2(se)$	$\beta^3(se)$	$\beta^4(se)$	$\beta^5(se)$	α
-3.77	-2.18	-0.60	0.06	1.20	1.26
-3.64	-2.29	-1.22	-0.52	0.85	1.20
-2.55	-1.37	-0.40	0.26	1.50	1.83
-2.20	-0.97	0.03	0.73	2.01	1.26
-2.06	-1.22	-0.41	0.21	1.07	2.65
-2.08	-1.27	-0.41	0.16	1.15	2.92
-2.15	-1.37	-0.64	0.00	1.09	2.00
-2.42	-1.23	-0.43	0.23	1.21	2.32
-2.37	-0.96	-0.07	0.58	1.71	1.59
-3.04	-2.00	-1.27	-0.80	-0.07	1.61
-1.33	-0.38	0.45	1.12	2.06	1.50
-1.98	-1.00	-0.20	0.32	1.36	2.51
-2.50	-1.50	-0.66	-0.14	0.88	1.16

A.2.2 TCQ IRT Analyses

Table A.15 shows the estimated coefficients for the two parameter GRM fit to the conventional items of the TCQ in Sample One.

Table A.9: Coefficient Estimates for MAAS, Two Parameter Graded Response Model, Split C

$\beta^1(se)$	$\beta^2(se)$	$\beta^3(se)$	$\beta^4(se)$	$\beta^5(se)$	α
-3.77	-2.31	-0.65	0.14	1.51	1.00
-3.62	-2.37	-1.08	-0.19	1.00	1.08
-2.41	-1.18	-0.28	0.48	1.69	1.75
-2.21	-0.95	0.08	0.94	2.22	1.24
-2.34	-1.29	-0.41	0.23	1.22	2.49
-2.40	-1.29	-0.40	0.24	1.47	2.71
-2.79	-1.54	-0.69	0.23	1.65	1.66
-2.55	-1.42	-0.55	0.18	1.36	2.26
-3.47	-2.30	-1.45	-0.77	0.08	1.29
-1.51	-0.31	0.68	1.36	2.48	1.26
-2.14	-1.19	-0.31	0.43	1.45	2.73

Table A.10: Two Factor Solution, LOT-R, Split B

Pessimism	Optimism	Communalities
-0.07	0.79	0.56
0.74	-0.03	0.53
0.08	0.67	0.52
0.85	-0.04	0.68
0.68	0.08	0.53
0.42	0.43	0.57

Table A.11: Two Factor Solution, LOT-R, Split C

Pessimism	Optimism	Communalities
-0.1	0.71	0.41
0.87	-0.03	0.71
0.03	0.72	0.55
0.86	0	0.75
0.61	0.18	0.58
0.24	0.59	0.64

Table A.12: Coefficient Estimates for Two Parameter Graded Response Model, LOT-R, Split B

$\beta^1(se)$	$\beta^2(se)$	$\beta^3(se)$	$\beta^4(se)$	α
-1.86	-0.23	0.88	2.16	1.15
-1.25	-0.22	0.49	1.66	2.05
-1.17	0.31	1.01	2.22	1.41
-1.11	-0.14	0.64	1.45	2.58
-1.13	-0.09	0.64	1.52	2.30
-0.83	0.25	1.16	2.09	2.43

Table A.13: Coefficient Estimates for LOT-R, Two Parameter Graded Response Model

$\beta^1(se)$	$\beta^2(se)$	$\beta^3(se)$	$\beta^4(se)$	α
-1.10	-0.27	0.39	1.51	3.11
-1.05	0.27	0.88	2.13	1.55
-1.09	-0.18	0.43	1.39	3.64
-1.18	-0.08	0.56	1.66	2.58
-0.72	0.30	0.96	2.02	2.29

Table A.14: Average of Five Factor Solutions over Splits B and C

	Pillinj	Hom	Rei	Cream	Acu	Communalities
1	0.6	0.11	-0.15	0.13	-0.11	0.52
2	0.61	0.13	0.01	-0.01	-0.02	0.53
3	0.56	0.08	-0.01	0.11	-0.04	0.48
4	0.72	0.11	0.14	-0.05	0.01	0.66
5	0.72	0.16	0.09	-0.01	-0.1	0.69
6	0.74	0.15	0.13	-0.08	-0.05	0.71
7	0.04	0.33	-0.06	0.48	0.03	0.68
8	-0.02	0.46	0.02	0.43	0.04	0.84
9	0	0.43	-0.07	0.47	0.03	0.7
10	-0.02	0.5	0	0.47	-0.01	0.89
11	0.04	0.49	0.05	0.44	-0.03	0.89
12	0.02	0.47	0.01	0.49	-0.01	0.92
13	0.62	-0.08	-0.13	0.1	-0.01	0.43
14	0.83	0	-0.1	0.05	0.11	0.75
15	0.68	-0.06	-0.05	0.01	0	0.47
16	0.87	0.02	-0.1	0.02	0.09	0.8
17	0.89	-0.01	-0.07	0.01	0.09	0.83
18	0.9	-0.03	-0.02	0	0.04	0.79
19	-0.08	0	0	0.12	0.82	0.76
20	0.04	0.01	0.01	0	0.93	0.87
21	0.01	-0.04	0.05	0.1	0.81	0.78
22	0.01	0.03	0.06	-0.05	0.93	0.91
23	0.04	0	0.03	-0.01	0.94	0.92
24	0	0.04	0.06	0	0.92	0.93
25	-0.01	0.44	0.03	0.46	0.05	0.86
26	-0.01	0.5	0	0.45	0.01	0.93
27	-0.01	0.43	0.05	0.47	0.02	0.86
28	0.01	0.52	0	0.45	0.02	0.96
29	0.01	0.49	0.06	0.46	0.01	0.96
30	0.01	0.49	0.05	0.47	-0.01	0.97
31	-0.05	0.04	0.81	0.12	0.03	0.87
32	0	0.02	0.92	0.03	0.02	0.95
33	0.03	0	0.88	0.05	0.03	0.89
34	0	0.05	0.92	-0.02	0.05	0.93
35	0	0.01	0.95	0	0.05	0.96
36	0.01	0.03	0.93	0	0.04	0.95

Table A.15: Coefficients for TCQ 1 Conventional Two Parameter Graded Response Model

$\beta^1(se)$	$\beta^2(se)$	$\beta^3(se)$	$\beta^4(se)$	α
-2.27	-1.09	-0.68	1.47	1.59
-1.87	-0.89	0.03	1.93	1.38
-2.26	-1.34	-0.66	1.67	1.76
-2.46	-1.10	-0.36	1.76	1.63
-2.20	-1.04	-0.34	1.79	1.79
-2.32	-0.39	0.47	3.07	1.03
-2.45	-0.59	0.47	3.73	1.10
-1.84	-0.20	1.05	3.65	1.15
-2.14	-0.59	0.56	3.98	1.23
-2.01	-0.30	0.57	3.28	1.33
-2.14	-0.53	0.52	3.39	1.33
-1.73	-0.97	-0.59	0.93	2.28
-1.92	-1.41	-0.74	0.70	2.98
-1.64	-0.82	0.12	1.34	1.97
-1.79	-1.15	-0.63	0.70	3.74
-1.55	-1.03	-0.48	0.72	4.04
-1.62	-1.08	-0.37	0.78	4.26

Appendix B

Appendix II: Qualitative Analysis

B.1 Introduction

Qualitative analysis was an essential part of this project. It gave insight and data into the development of the IAT's used in the final part of the research project. A thematic analysis (Braun & Clarke, 2006) of the interviews conducted was carried out to develop themes both for the repertory grid and for the IAT.

Qualitative research typically relates to the analysis of interviews and other texts derived from people. It differs fundamentally from quantitative analysis in that it aims for a deep understanding of particular individuals, while quantitative analysis aims for a broad understanding of the sample as a whole.

The biggest problem with qualitative analysis is that it cannot scale to the level of large scale surveys, as it requires significant amounts of researcher time per participant while quantitative surveys have a cost of development in time, but the marginal cost of administering the survey to a new participant is essentially zero (assuming distribution over the internet).

The issue of reflexivity is crucial to qualitative research (and also appears in quantitative research, though rarely as openly) (Rosenthal, 1967, 1969).

Reflexivity refers to the impact of the researcher's prior conceptions and approaches have on the course of the interviews (Finlay, 2002).

This is extremely obvious in the choice of the major questions to be asked in the interviews, but it can occur in subtle ways during the interviews also (for example in the use of language by the interviewer) and in the quality of communication or rapport experienced by the researcher in the course of the interview. Reflexivity is also critical during the analysis, as the researcher must be aware of their own biases and ensure that this affects the analysis as little as possible, or at least report where the problems arose for them.

B.2 Analysis of Health Construct Interviews

B.2.1 Methodology

Eight interviews were conducted in a semi-structured format. The participants in the interviews included 3 alternative/complementary therapists, 2 GP's and 3 students with experience of health-care, both western and complementary.

The questions which were asked were as follows:

1. What does the term health mean to you?
2. Have you had any health problems in the past that you would like to discuss?

3. How do you believe that health can be promoted and sustained?
4. What kinds of treatments do you feel are effective?
5. How are they effective?
6. What has been your experience with the official medical sector?
7. What has been your experience with the complementary medicine sector?
8. When you are sick, what types of help or treatment do you find most useful?
9. What do you think are the major causes of sickness (for you or others)?
10. What people or qualities do you associate with health?

The Interviews were conducted between October 2009 and January 2010.

B.2.2 Analysis

The interviews were first coded using an inductive coding procedure. In this, the codes were developed from the interview transcripts themselves, allowing the participants to define the nature of the analysis (in some sense). Following completion of the interviews and first coding, the interviews were then re-coded using the entire database of codes which had been developed throughout the analysis procedure. During this phase redundant codes were also removed. Following the coding procedure, codes were combined into thematic units where appropriate. For example, the many codes relating to health were combined into an overall health code.

This analysis will look at the material collected and analysed in the following ways:

1. Analysis of themes across the entire sample
2. How this relates to the core subject of the thesis

B.3 Methodology

B.3.1 Thematic Analysis

The thematic analysis's primary purpose was to look for common patterns in the conceptualisation of health, sickness and treatment. As such, it was felt that the best approach would be to develop the codes from the transcripts themselves. This is called an inductive approach to coding (Haberman, 1979). Following transcription, each interview was coded line by line by the primary researcher, and codes were developed throughout this process.

After all the interviews had been transcribed, the codes were pruned and amalgamated to reduce redundancy, and this process was repeated. This second coding lead to a number of new codes and insights which had been missed the first time, and the text was again coded for a third time following the development of these new codes.

Then, the document was coded a fourth time, but on this run through the aim was to look at higher level patterns that emerged from the text. Following this coding procedure, a process of chunking of codes was carried out. This involved looking at how codes fit together and grouping them under a number of thematic headings. The original texts and recordings were referred back to at this point to ensure that the themes were representative of the original data, and finally the themes were written up to record the results of this exercise.

B.3.2 Participants

B.3.2.1 Alternative/Complementary Practitioners

Three participants were interviewed as part of this group. One was a shiatsu and Reiki practitioner, one was a spiritual healer and the third was an acupuncturist. They were all recruited from the Cork area following email contact through an alternative therapy website.

B.3.2.2 Doctors

The two doctors were both General Practitioners recruited through the UCC school of medicine.

B.3.2.3 Students

The three students were all recruited by asking participants in the first TCQ Sample (see Chapter 5).

B.3.3 Overall theme

In these interviews, a number of important themes emerged from analysis of these interviews. These are noted here and further developed throughout the test. These major themes were as follows:

1. Health – internal or externally determined. This theme seemed to stratify respondents into groups. Some respondents (alternative therapists mostly)

seemed to look at the determinants of health as being mostly cognitive and emotional, while others (GP's mostly) seemed to look at more external determinants of health, such as pathogens and social status. The ordinary people interviewed in this research project tended to incorporate elements from both of these perspectives;

2. Doctors as prescribers, alternative therapists as facilitators – the alternative therapists tend to talk about facilitating people towards health, while the doctors tend to talk about curing people;
3. Energy versus biology – doctors tend to talk about biology, alternative therapists tend to talk about energy;
4. Symptoms versus root cause.

B.4 Results

The results section for this chapter focuses on the categories which were coded from the interviews, with some of the original codes reported in the appropriate place.

B.4.1 Health as balance

This code occurred quite frequently, but only in the transcripts of the alternative therapists. Some examples and discussions follow below.

“So these are the processes in the mind, so again, health in respect to that, to emotions, mental energy very much keeping on a balanced level – Alternative Therapist 3”

We see here that the participants believes balance to be key to all health, on an emotional, physical and mental level. We can also note the focus on energy, which will be examined further below. We can also see here a focus on processes, that these elements are constantly moving, which is borne out in the quote below.

“The same thing with the emotions, keeping them moving, keeping them flowing, ammmm – Alternative Therapist 3”

Again, we can see the focus on movement, that emotional expressions should be facilitated, and we can imply conversely that stagnation of emotions is unhealthy.

This is interesting in that there is much psychological evidence that suggests that secrets and unexpressed emotions can link into physical health problems.

“Same thing, the imbalance- Of emotions, imbalance of mental thinking imbalance of sleep not sleep to much sleep ahhh and then what you’re putting into the body obviously, and also what you’re putting into the body if you’re sitting in front of a computer for too long, if you’re sitting in front of a TV too long, you’re putting too much electromagnetic energy on the mobile phone for too long so generally, ammm so you’ve cardio-vascular health we need balance there—
Alternative Therapist 3”

We can see from the above quote that while health is conceived of as balance, sickness is thought to result from either too much or too little of a particular substance or process. We can note the focus on both mental and physical imbalance here, this linking of levels of existence is a theme that pervades the entire transcripts, and is called out throughout. The interesting point about this theme is the way in which it provides a framework through which the participants can make sense of their own and others health.

B.4.2 Energy Model of Health

Another code which occurs throughout the transcripts of the alternative therapists is this notion of *elan vital*, or life force. This is one of those ideas with a strong resonance throughout human history, as multiple cultures appear to have developed in independently. This was looked at in more detail using excerpts from the transcripts below.

“how do I think they work? Well....from doing different treatments....like shiatsu, so which works on the energetic plane – Alternative Therapist 1”

We note here, that when the respondent was asked how some of the treatments used by them worked, they focus immediately on shiatsu (a form of Japanese massage) which is the treatment they use most often in their professional practice. We can also see that they construct this treatment as operating on an energetic plane, which presumably means that it effects the person on another level from the physical. This theme continues throughout the transcripts of the alternative therapists, and appears to be the explanatory model utilised by the majority of them.

“you actually work energy channels that ammm...remind....the person to put energy in certain places. Yeah, like I would press certain acupuncture points, work certain energy channels...- Alternative Therapist 1”

As we can see from this excerpt, the practitioner seems to link this energy body to the physical body, such that changes in physical pressure on certain points, suggesting that these two parts of a person are interlinked and occupy the same space.

“it is too easy to say mind body soul but roughly its pretty much that, physical energy, taking the energy from your food and translating that into nutrition into energy and accommodating that and then expelling energy and expelling waste products efficiently that’s the physical side of things energy then – Alternative Therapist 3”

However, in this excerpt we have a very different interpretation of energy from this respondent. The energy they describe links up exactly with what we would consider energy to be (the breakdown of food into sugars), so this definition appears to differ quite substantially from the one given by Alternative Therapist 1 above, which suggests some conceptual confusion. Further excerpts will probably make this clear, however.

“There’s emmm very common ahhh which my professor used to say – if ahhh if you want to do exams, if you’re studying, eat well. He says the spleen produces your blood your red blood cells – Alternative Therapist 3”

This extract is interesting, in that the energy model is extended here to apply to mental efforts also. We can also see the linkages between physical and mental health here. Again, we see examples of organs in the body being associated with particular states and traits of the mind.

“ In France they use digestifs, and apertifs, and thats that’s the same thing, you know the same as preparing the energetics to assimilate....transform and bring forth the energy from that food very very carefully, and if you do that, then you’re going to have Chi and if you’ve Chi the the character for Chi in Chinese medicine, or sorry in Chinese language ammm the I’ll draw for you is is similar to that and basically what it symbolising is a fire, a pot with rice in the pot and steam rising from it so the pot is sitting over the fire cooking the soup – Alternative Therapist 3”

Again here we see the linkage of physical nutrition and energy. This respondent links the physical act of eating to the life force describes as Chi, suggesting that this Chi terminology may be a descriptor for the sugars and other products needed to keep humans alive. Again, we see the notion that food also gives us this energy, which is separate from the physical value of food, and that this is the energy which can impact our health and the lack of which causes us to become sick.

“So again, when you say how do I think it works thats how I think it works as well manipulating energy P: ammm, the energy has to be there first day or first place, so someone is very depleted then often I have to make sure I that we get them to a place where they can that I can manipulate energy because if it’s not there, I’m only going to deplete them much more – Alternative Therapist No 3”

Here we see a construction of this energy, or Chi in that it can become depleted, and that the task of the therapist is to increase this energy in order to allow their methods to work on the patient.

“am manipulating pockets of either static or depleted for replete excessively replete some energy in the body and ahhh.... I: that effects change? P: Yeah, exactly, yeah, yeah. So thats I suppose my limited way of understanding it and the rest then is down to patient cooperation, really, go and do your rest, we’ve told you enough here – Alternative Therapist 3”

In this excerpt we can see that the problem is constructed as one of imbalance, either there is too much energy, or too little and that the role of the therapist is to balance these energy pockets across the person to ensure that health returns.

B.4.3 Emotional Health

This code was predominantly used by the alternative therapists, but also appeared a number of times in one of the GP transcripts.

“P: Ammm...so....a non-stressful lifestyle you know? And a ...fun, laughter, that you do things that you like to do. - Alternative Therapist 1”

“o its more preventative, in its focus more and its all about balance being in P: that’s all the physical side ... P: and then I try to definitely you know, meditate,

... P: and ehhhh, focus on my emotions and my thoughts – Alternative Therapist 1”

We can see from these quotes that for many of the alternative therapists, emotional well being is an important factor in their construction of health. They tend to construct health as being situated in a context whereby emotional well being is extremely important and that it forms an integral part of health. This links in with the research suggesting that optimism is associated with health, something which will be discussed later and which appears to be a construction made by almost all of the participants in this study.

“P: Most causes are the emotional- P:and so the body is really reacting and that P: So thats – thats why we say mind and body, you know that kind of a balance – Alternative Therapist 2”

For this respondent, the emotions are the primary driver of sickness or health. They appear to construct health as being something primarily determined by one’s emotional outlook, rather than resulting from pathogens or the environment. Its interesting that this is exact opposite opinion offered by the GP’s who were included in this study.

“emotional energy has to be very well balanced we have to breathe our emotions, we cannot suppress them am they have to (pause) be (pause) tuned in and out of the body very carefully. - Alternative Therapist 3”

As we can see here, the construction of the importance of emotions runs throughout all of the transcripts of the alternative therapists. This respondent links the emotional expression with the breathing, that they have to be taken in and let out very regularly. They seem to construct them as something which can have great physical impact, and should be treated with respect. Its also interesting to note the strong use of movement-related metaphors, something characteristic of this respondent is that health is constructed as a process.

“ahhh it it if used in itself it may not answer the whole problem because people come not only with a condition, they come with with a condition attached to their own bodies its got feelings, its got thoughts preconceptions, its got worries, etcetera- - Doctor 2”

We can see from this quote that this respondent possesses a radically different construction of emotions and their relationship to health. For them, the body is the primary part, and they construct this using depersonalising language (it) the body is the important part, but it has all of these messy emotions and thoughts attached to it which complicate matters. The contrasting quotes here nicely illustrate the differences in construction between the alternative therapists and the general practitioners.

B.4.4 Personal Effort

Again, this code appeared mostly in the transcripts of the alternative therapists, although it also appeared in the transcripts of one of the doctors.

“reflex points, If a person is not willing to work on themselves and to get better, you can do whatever you want, and they- Alternative Therapist 1”

“so it requires their willingness to take part in the procedure aswell - Alternative Therapist 1”

These quotes serve to illustrate a point made again and again by the alternative therapists and to a lesser extent by doctors. That point is that people are ultimately responsible for their own health, regardless of what help they get from a health-care professional.

“Yeah, but again it will need a change in lifestyle, change in diet, a willingness for the person to actually change – Alternative Therapist 1”

Here we can see a restatement of the points made above, that while the practitioner can give advice, the person must ultimately act upon it, and if they are not willing, then they will not change.

“Yeah, I don’t fix people, they have to fix themselves then with support and inspiration from me – Alternative Therapist 2”

This is perhaps the most telling quote in this section. This respondent constructs patients as being totally responsible for their own health, that he merely acts as a

facilitator and allows them to heal themselves. This is a very empowering view of the healing process, and contrasts quite strongly with the more paternal constructions of the doctors. Its also interesting to note that the word used is “fix”, which seems to suggest a sense of the body as a machine.

“how much they’ll try to put practices that could be suggested in the clinic into into their lifestyle ammm, you can also I think tell if if they’re . . . you know you point something out to someone, that they can make this subtle change that will make a huge difference to their bowel habits P:Suggest something for them to eat- P: almost on hearing it, I could there could almost be a placebo effect would not be surprised when they come back in two weeks and tell me that their bowel habits have completely changed but I would often almost be able to tell with a patient who’s going to do that or who’s going to be..... a little bit more open to that or ammm.....how else do I think it works. I I I think a lot of the time with any form of healing I mean, - Alternative Therapist 3”

Again, we have a construction here that points out the differences between people’s response to their health. Some other respondents suggest that health is not valued by many, and this respondent argues that some patients will do the work required to get better, while others will not, and that it is this factor which accounts for broadly divergent outcomes.

Even though explicitly called out, the sense that some patients will get better (and the emotions and behaviours elicited by this sense) may well be a factor associated to better recovery.

“because people think that everything going right comes without any effort and in actual fact, to make things right requires huge effort – Doctor 1”

We have this quote from one of the GP’s here, and it supports what the therapists have been saying. One can say that health is being constructed as something which needs to be valued by people if they are to enjoy good health. Again, it can be seen how the doctors tend to use more impersonal language, perhaps an artifact of their training.

B.4.5 Body Knowledge of Health

Again, this code appeared only in the transcripts of the alternative therapists, and only in two of three of those. Below are some relevant examples with some commentaries on their constructions of health.

“Because you should more talk to your own wisdom but if you really sick of course you have to go to the doctor – Alternative Therapist 2”

This is an interesting quote, in that it seems to construct an awareness on the part of the person as to the causes and nature of their sickness. The body appears imbued with understanding of the problems facing the person, and this can be consulted. It seems however, that this is only a first resort, as the quote continues to state that serious problems should be dealt with by a doctor.

“By paying attention to what is (pause) going on with yourself. We’ve simple basic needs. We need to sleep and we need to eat. Basically, those are the two most important we also need to breathe. Otherwise we can – we can name ammm – Alternative Therapist 3”

Here again there is a focus on internal knowledge. This time, however, it appears constructed in terms of awareness of the body’s needs at a more basic level, that of sleeping and eating. This contrasts with the unspecified wisdom referenced in the first quote. However, the interesting emphasis occurs on the word breathe, which does not seem to just refer to the physical, mechanical act but to something more metaphorical. This comes across from the repetition (and indeed, the voice transcript shows an emphasis at this point).

“P:And to reassure them you know that, everything is OK with them in in some fashion as well and its sustained by (pause) listening to yourself, and listening to nature – Alternative Therapist 3”

Here, from the same respondent we have an expansion of the previous statement. Where first listening to yourself was important, now we have a focus on listening to nature also. While we can construct this as listening to the natural world, this construction could prove quite problematic as there are many different ‘natures’, and this quote does not allow us to distinguish between them.

B.5 Construction of Doctors

One of the major themes that emerged from all of the interviews was the nature of doctors, even apart from a more general conception of health practitioners, all the participants focused on the role of doctors in society, and their positive and negative impacts.

One of the ordinary participants focused on the relationship between Doctors and Alternative therapists, with this excerpt “I think if you have like serious [pause] health problem, you will need both”. This stresses and indeed this participant stressed throughout the interview, that doctors and alternative therapists should be viewed as complementary rather than opposing.

“she mentioned this to her GP who (pause) became very angry and asked me to phone him immediately P:and this type of thing doesn’t happen very often. I rang him, and he was quite set, saying I don’t want this woman to have more and more treatments, you know and I want I want her to find the right treatment. And I said, so do I, and this is it, and this is what I deem most good for her at the moment and he asked about the treatment and I tried to explain a little – in simple language – and the end of the conversation he was Ok, let’s do this.”

This quotation came from Alternative therapist 3, and describes the relationship between doctors and alternative therapists. Note that he constructs the two as in a collaborative relationship rather than competing, and shows that they were able to find a good balance between both of their particular forms of treatment.

It is interesting in that one of the doctors focused on their role as agents of social change, particularly in this excerpt

I think as doctors we have a job maybe as advocates you know to point out the issues, and point out the we may not be have the solutions but we should be able to sort of say these are part of the problems – these are the problems, and these are some of the determinants and really these need sorting out

This quotation focuses on the role of doctors in society rather than their role as a individual health practitioner, making the point that they have a responsibility to their patients to focus on the problems that are actually affecting their abilities to live healthy lives.

Interestingly, many of the alternative practitioners did not subscribe to this viewpoint of doctors, instead regarding them as prescribers

Ehhh, a lot of doctors just ah rely on on on on eh their pharmaceutical companies

The implication here seems to be that the commercial relationships between doctors and pharmaceutical companies get in the way of healing.

However, this point was actually raised by one of the doctors themselves, in a slightly different context:

I think doctors actually should have a ethical obligation to say enough is enough you know and I get actually quite frustrated when I get sent a patient that I actually feel has nothing wrong with them - Doctor 1

Here we can see a doctor's frustration at the way in which they are expected to dispense medicines to patients even when they feel there is nothing wrong with them. It is worth noting that the doctors frame this as the patients demands, while to the alternative therapists, this is a relationship in which the doctors have power over the patients. Its also interesting to note that there is a sense in which this represents a judgment of the patient and their perception of illness, which is something that the alternative therapists in this sample did not tend to talk about very often.

The second GP also expressed frustration with this state of affairs, saying that

even though you try to spend as much time as you can in terms of health prevention such as immunisation or health promotion by giving advice to people on healthy diets exercise giving up smoking etcetera we do tend to spend most of our time reaching for the pen, prescribing - Doctor 2

The sense is that they would like to be able to focus more on the health promotion effects that would actually change the persons state of health, but are corraled by the system into prescribing, as that is something that can be done within the confines of the 30 minute doctors appointment.

Also, one of the students expressed some annoyance with elements of the system:

well as I say like, just that kind of stuff ammm like plenty of times i've been sick and gone to a doctor and he's given antibiotic and i've taken it and gone thats not helping me at all -Student 2

The notion is that antibiotics are a default option for many doctors and patients, and for some of the respondents, this is not actually particularly useful.

not unless you're actually particularly sick like, I think that an awful lot of the time people tend to be a small bit sick and go off to a doctor and get their antibiotics and go this didnt help at all, whereas there was probably no need to go to a doctor in the first place -Student 2

The quotation above referenced another student's answer to the question how often do you go to the doctor. Note that given the age (early twenties) and social status (student) of this respondent this is perhaps not a surprising response. The theme that comes through here is that doctors can be associated with unnecessary treatments and methods, and that this is a conception shared by all three groups in the sample.

Another theme that emerged around doctors was that doctors focused more on illness.

The extract (from Doctor 1) illustrates the point:

I suppose sometimes health and ill health, you know, there's not different sides of the same coin P: but I suppose we're more set up in our training to deal with ill-health as a concept// I:yeah P:- than health and particularly how the health services are structured

The participant notes that ill-health gets far more attention than does health, perhaps linking back to the idea that ill-health is a breaking down of something that is normal (i.e. health).

Another quote from the other Doctor (No 2) supports this reading:

we'd only get so far, and I think that we essentially would be more into the disease ahhh ammm diagnosis and cure-

It is interesting to note that the doctor first says that they are more into disease, but then shifts their words to diagnosis and care, displaying the focus on ill-health referenced above. This also relates to the notion, expressed by some of the other respondents, that doctors tend to focus on problems.

Perhaps an interesting theme which emerged from the interviews with ordinary people and alternative therapists was that doctors were not that important. The interesting point about this is that by making the point, they were in fact reinforcing the notion of doctors as important.

So just to have a regular checkup and stuff checkup and ah, well everything was fine but yeah I don't really know so much about doctors. I think doctors are great if you need them but if you don't need them you shouldn't go to them.
-Alternative Therapist 1

The respondent here proudly proclaims their ignorance of doctors, and introduces some tautologies in that doctors are only necessary if you “need” them, but does not define what exactly it is to need doctors.

not unless you're actually particularly sick like, I think that an awful lot of the time people tend to be a small bit sick and go off to a doctor and get their antibiotics and go this didnt help at all, whereas there was probably no need to go to a doctor in the first place

Here the respondent focuses on doctors only being important when you are really sick, and argues that people tend to go to doctors for issues which they would not consider important enough.

The next theme around doctors came from Alternative Therapist 3, where they recall an incident where doctors came to see them.

initially came for pain a third came for tiredness very interestingly they all asked to come after-hours I: Because they didnt want to be seen P: By their patients, they know my system and that I do one patient per hour and I don't do two or three rooms like a lot of people do and they were fine with that ammm (pause) one a patient (pause) didn't want necessarily anything other than the main complaint which I think was stress ammm but I asked could I treat for other matters

The interesting part here is that he constructs this incident as something out of the ordinary (whereas the converse is broadly accepted by all participants) and frames it as a series of differences between the doctor's typical practice versus that of the alternative practitioner. Note also that the doctor only wanted treatment for the main presenting symptom, rather than a more holistic process which the alternative therapist would have preferred.

Another intriguing theme that came through from one of the doctors was the following:

the herbalists, who have almost taken on a bio-medical model then a bio-medical sort of way of acting they see patients they diagnose they have a pharmacy of herbs - Doctor 2

The participant was responding to a question around whether or not they had any experience with alternative medicines and treatments. Note how the “good”

alternative therapists are defined by their adherence to the bio-medical model, and note that the other salient features are diagnosis and pharmacies.

However, note the contrast here with how one of the alternative therapists describes their relationship with doctors.

I've become much more in happier terms today with it because of with the fertility side of things I need to rely on blood results a lot. I need people to go and have their hormones checked and and something like sperm analysis, I need, I need medical science completely for this these reasons. - Alternative Therapist 3

This respondent regards their relationship with medical science as non-adversarial, but rather complementary. The contrast with the attitude of the doctor is quite striking. There is a slight hedge in that the respondent suggests that they only need medical science for this part of their practice, and perhaps implies that it is not useful for anything else.

they just prescribe medication that has a lot of side effects and they don't really spend much time with you and you go to the doctor there's very few doctors which actually spend time and actually talk to people and really see where the problem comes from- Alternative Therapist 1

Here, the conception of doctors is as seen earlier, distant prescribers who do not actually spend time with a person. The implied contrast with their own practice is apparent. Note the notion that the respondent focuses on the root causes, while doctors focus on the symptoms.

Even some of the Doctor (No 1) respondents argue in this point:

I would disagree with what my colleagues might do or I've occasionally been shocked ammm by what I would perceive to be a lack of care,

Here it is constructed more in terms of professional disagreement, and the occasional problem, whereas the alternative therapists are more forthright about this (but less so about the problems with their own profession).

And from one of the Doctor participants:

ammm, a bit, I'm not trained in them but obviously I prescribe St John's wort - Doctor 1

Note the unassuming tone, a bit, expressions of ignorance but the use of obviously to describe the practice of prescribing St John's wort, as though it were the most normal thing in the world.

they're just trying to ammm [pause] you know [pause] cure the symptoms rather than the origin of the amm, imbalances -Alternative Therapist 1.

Again, note the contrast, which is that therapists focus on the origin of problems, while doctors treat the symptoms.

Finally, a doctor gives some credence to this argument, however attributing the blame for this situation elsewhere.

then some people are quite happy with a model of you know diagnose me, give me the treatment -Doctor 2

Note here that there is an acceptance that many doctors just treat symptoms, but the blame is placed on the patient for this, who seem to be regarded by the doctor as giving up their own responsibility for health when they see the doctor.

B.6 Diet

The next step in the analysis was looking at the codes and themes related to diet.

B.6.0.1 Food as medicine

An example of this code is below:

"in China, they have even the most common person on the street he's a peasant working in the fields are born and have and inherent knowledge of how to eat in accordance with nature right in their body for a cold condition they know to eat something warm corn soup with ginger or rice casserole, meat and if they've a very very hot condition they know to eat know to eat, you know, cool foods and its very much putting nature into the body looking at external and balancing between the two like I was mentioning a while ago- Alternative Therapist No. 3"

As we can see in from this code, the conception (shared by folk healers across the world) regarding hot and cold foods exists here in his description of health. We can also see how he constructs this as a complement to Western styles of medicine, that it is common knowledge in other, non-traditional societies (where his method is drawn from) but is ignored in our prevalent system of medicine here in the west.

Another example (from the same participant) follows below:

“I personally do like herbal medicine and I think that its its as close to food as you can get a herb generally grows either above or below ground the roots are usually stronger versions of it so just knowing how herbs work I will use preventative medicines and supplements to protect say if I find working too late head full of information and details, just go for a walk and just try and clear”

Here the participant is talking about what he does when he is sick, and we see how he legitimises herbal medicines by linking them with food, which he has previously described as the best method for promotion of health.

“good food — healthy food organic food”. Alternative Therapist No. 1

We can see here (this participant tended to give very short answers to questions) that the emphasis on food is common to the alternative therapists here. We can also note the linkage of health food, organic food as possibly indicating some disquiet with the technocratic nature of herbal medicine.

B.6.1 Diet as Fashion

So, we’ve a habit of here in the west of -again, going back to diet – people will eat any food at any time and ammm ammm menus in this country have become cosmopolitan palatable ammm you know, fashion as a menu read, or if you’re choosing off a menu be choosing for what the needs are Alternative Therapist 3.

Here the notion is of food as a lifestyle accessory for many people, contrasted to the actual needs. Some of the thoughts of this therapist on useful foods are discussed below.

Nor is it sustaining your health to eat cold salads all the way through winter, and yet people will do that Alternative Therapist 3

Again, there is the notion here that certain foods make much more sense at different times of the year, and a certain sense of the disregard of people for what is obvious to the practitioner

B.6.2 Diet as Place Based

So not everyone wants to eat hot spicy foods because you know, hot spicy foods are designed for countries like India you know, Bangladesh and Southern China the hottest spices in the world where its incredibly hot and they eat these foods to open pores, to sweat and to purge- P:to balance themselves with the he-weather whereas we eat them here because its trendy to eat something with chili. (Alternative Therapist 3)

Here again, the notion is developed of certain foods being appropriate for certain places but not others, and again the idea of “fashion” is developed when discussing the food habits of their patients.

B.6.3 Nutrition

The final section in the food themes constructed was one of nutrition.

healthy food organic food Alternative Therapist 1

Again, here the notion is developed that healthy food is organic food. This harks back to much of the focus of the alternative therapists of perceived naturalness and sensitivity to the environment.

Yeah, but again it will need a change in lifestyle, change in diet, a willingness for the person to actually change Alternative Therapist 1

Here again, the notion is developed that the change in diet must be accompanied by further changes, that the levels of change are in some way connected.

would support myself then with ehhhh, supplements and ehhhh, eh hh, herbs, homeopathy, [pause] and a diet Alternative Therapist 1

Here, the notion is that supplements are important (as distinct from pharmaceuticals), and that diet is one part of an overarching system.

they say most of the time, you know, what you take too much sugar your body also reacts on that Alternative Therapist 3

Here, we have another alternative therapist is describing the opinions they have regarding nutrition (and especially sugar). Note that the speaker distances themselves from the utterance, by suggesting that “they” say this rather than expressing his or her own beliefs directly.

taking the energy from your food and translating that into nutrition into energy and accommodating that and then expelling energy and expelling waste products efficiently that’s the physical side of things energy then Alternative Therapist 3

Here we have a succinct description of one of the therapist’s point of view of physicality. Note that this is all framed in terms of energy, as are the therapists own treatments. While the description is accurate, the contention is that this physical level is just one amongst many.

You need physical, you need food so food is the essence Alternative Therapist 3

Here again we see that the speaker focuses on food and diet as being extremely important for health, but this is also framed within the greater context of essence and energy.

right in their body for a cold condition they know to eat something warm corn soup with ginger or rice casserole, meat and if they’ve a very very hot condition they know to eat know to eat, you know, cool foods and its very much putting nature into the body looking at external and balancing between the two like I was mentioning a while ago

Here we have the juxtaposition of the notion of hot and cold foods in the speaker’s words. The idea is that there is a balance between hot and cold (or dry and moist)

which is an idea which occurs in the philosophies of many older cultures, such as the Greeks, as recorded by Aristotle.

health promoted and sustained as I said dietary is obviously the No 1 Alternative
Therapist 1

Here again we see that this speaker focuses on diet as the most important driver of health.

the most important thing that they can leave the room with is an understanding
of how to eat over the next month to maintain their health or keep it well.
Alternative Therapist 3

Here we see the therapists answer to how health can be promoted and sustained, which is that diet is the most important factor in maintaining a healthy standard of living.

But really my (pause) ammm strongest push on it would be dietary means-

what I would do is I will watch for the warning symptoms, cold and cough a lot
of people in very close perimeter ammm, you know, dietary side of things, if I
notice anything that's a little out of kilter I'll try and address it through diet try
and you know a-ahhhh you know yourself very well, ammm in the system of
medicine cos you do so much detailed consultations with people it difficult not to
think in the same way about yourself - Alternative Therapist 3

Here we see his conception of how his treatments work, and how his health is maintained. Note that he frames this as a continuous process, paying attention to his somatic sensations and then remedying any deficiencies through diet. There is also an acceptance that what you do for your own health is strongly shaped by the treatments that you provide for others.

Apart from that I suppose Irish people in particular I think especially people in
college mightn't know ammm how to eat nutritious foods properly or how to they
mightn't get regular exercise I think they obviously have a good contributing
factor to that too like Student 1

Note that here the framing is that individuals lack knowledge about the best nutritional choices to make, rather than that it comes from any sense of “fashion”. Again, it is interesting that it is framed as Irish people in particular rather than people in general.

I would feel you know therapy or ammm there'd be you know certain foods that might help in the diet or what have you and then that can lead to positive mental health for example Student 2

Again, here the final quote refers to foods that are important for both physical and mental health, in that eating particular foods is framed as being somewhat causal for both physical and mental health. In this, the respondent is linking to some of the other themes around the linking of body and mind.

B.7 Models of Health

The next set of themes deal with the conception of health by different respondents.

B.7.1 Health as balance

Health as balance was a theme that emerged primarily from the interviews with alternative therapists, and to a less extent, from the interviews with students.

and I try to find out what ammmm, [pause] where I went off balance Alternative Therapist 1

Here, the respondent is talking about what the causes of sickness are. There is no mention of external forces here and the conception is that getting out of balance appears to be a primary driver for less than optimal health.

kind of things back in balance Alternative Therapist 1

Here, when talking about getting well, the same metaphor is used to describe this process. The difficulty here though, is figuring out exactly what kind of balance the respondent is talking about.

The words of another alternative therapist may prove useful here.

mind and body, you know that kind of a balance Alternative Therapist 2

Here the balance is between mind and body, harking back to the old saying about a healthy body in a healthy mind.

emotional energy has to be very well balanced Alternative Therapist 3.

Here we see the linking of the notion of balance to the notion of energy, which were two key themes which emerged from the interviews with the alternative therapists.

right in their body for a cold condition they know to eat something warm corn soup with ginger or rice casserole, meat and if they've a very very hot condition they know to eat know to eat, you know, cool foods and its very much putting nature into the body looking at external and balancing between the two like I was mentioning a while ago

Here again it can be seen that the respondent links the notion of balance back into the system of hot and cold foods and their impact on diet. This theme of balance seems to pervade the interviews with the alternative therapists.

then again maintaining health its all about balance, so you don't leave it too on one side cos you leave yourself too open ammm you know, there has to be balance within it. Alternative Therapist 3

The theme of balance comes through here, as even balance is regarded as something to be balanced.

So these are the processes in the mind, so again, health in respect to that, to emotions, mental energy very much keeping on a balanced level

Again we see that balance is to be achieved between emotions and mental thought, expressed in this language of energy (while using quite physical metaphors, level, for instance).

sickness are? P:Same thing, the imbalance- P:Of emotions, imbalance of mental thinking imbalance of sleep not sleep to much sleep ahhh and then what you're putting into the body obviously, and also what you're putting into the body if you're sitting in front of a computer for too long, if you're sitting in front of a TV too long, you're putting too much electromagnetic energy on the mobile phone for too long so generally, ammm so you've cardio-vascular health we need balance there- Alternative Therapist 3

Here again the focus is on sickness as being caused by a lack of balance, however, balance is construed in a much broader sense in that particular impacts of the environment can have large effects on health and need to be balanced in particular ways. For the respondents it seems like health is constructed by a process of figuring out what is needed to offset the effects of the environment.

B.7.2 Health As Harmony

So for example, when again the question what does health mean its being its it sounds very cliched but living in accordance with nature. Alternative Therapist 3

Again, the respondent constructs health in terms of being in balance (c.f. earlier quotes in the health as balance section), and living in accordance with nature, although nature is never defined in this statement.

in China, they have even the most common person on the street he's a peasant working in the fields are born and have an inherent knowledge of how to eat in accordance with nature Alternative Therapist 3

Here we have the typical conceit that exotic places have knowledge that is more true than those around us. The conception of China is perhaps idealised, but the theme is that we as Westerners are somehow disconnected from our "true nature".

:And to reassure them you know that, everything is OK with them in in some fashion as well and its sustained by (pause) listening to yourself, and listening to nature

This particular quote is interesting, in that it links the beliefs noted above with the practice of working with patients. Again, the conception is around being receptive to the internal and external signals of the world.

But in China they wouldn't do that Could be, you know, they'd order any person on the street would not eat salad during the winter cos they know its the food is cold energetically - Alternative Therapist 3

Here again the use of the Chinese as a foil to the people around the respondent, and the contrast to his current patients (see the Food as Fashion section). Note that the respondents starts with a positive statement, and yet continues with a declaration of what these people wouldn't do, suggesting that this is framed in a positive light.

Its kind of similar same thing that macrobiotics from Japan eating in accordance with nature, eating the foods that grow around you and you know work we're coming into winter now, so eating root veg is more – is much better idea than eating veg that grow above the ground because when the frost come, they will die anyway - Alternative Therapist 3

Here again, it can be seen that the notion is to be in harmony with the internal and external environment. The focus here is on food, but the general theme reappears again and again in excerpts from this participant.

B.7.3 Health as Social Issue

Yeah, I think it would help, but I also think even simpler things even going to the doctor you see is prioritising health P: and health is not a medical problem, health is a social issue you know Doctor 1

This quote emphasises the difficulty with health for many doctors, the profession focuses on solving problems (i.e. symptoms) but “health is not a problem”. The doctor constructs health as being far broader than sickness, as a matter for society rather than the doctor.

and health is wider than just a doctor its everybody- P: and society's problem I:- society's problem as it were

This quote reinforces the previous excerpt, in that health is framed as everybody's responsibility. To some extent, this is a tactic which shifts attention away from the responsibilities of the respondent, but it seems to be a relatively useful insight.

well I think first of all its something that is beyond the realm of the health practitioner Doctor 2

Here again, the other doctor respondent focuses on health as something beyond any practitioner. It is interesting that both doctors respond like this, in effect throwing up their hands at the whole matter.

so in a sense, we're there to pick up the pieces, ahhh while we have a role in health prevention and and actually disease prevention, health promotion ahhh there's an awful lot of other people right down to the individual themselves who have a huge role to play in that Doctor 2

The key message here is the first sentence — that doctors are there to pick up the pieces, rather than to ensure that they do not get broken in the first place. The respondent references the whole health care system before they consider the role of the individual. It is interesting to contrast this kind of approach to that of the alternative therapists.

Ill health is I: - a medical problem P: is a medical problem, but good health is much more ammm wider Doctor 1

Here the contrast between sickness and health is made explicit in that health is constructed as being much more encompassing than is ill-health. Again, the doctor states that this problem does not lie within the parlance of medicine.

there's positive attributes as well as the absence of the negatives

Again here the notion is made explicit — health is not merely the absence of sickness, but rather a thing unto itself.

B.7.4 Health is not a Problem

health is not.....you know, ill-health is a problem but good health is something that you need to invest your time in it, and unless you make that a priority for yourself you're not gonna I: you're not gonna have it P: you're not gonna do it-I:mmmm Doctor 1

Here again health is constructed as something that requires active work, in that it is not easy to get and needs to be made a “priority”. Again, note the response to the interviewer, the replacement of possessing health with performing actions that result in health.

and health is wider than just a doctor its everybody- I:yeah, yeah, yeah P: and society’s problem Doctor 1

Note the contrast here between the earlier statements around health - now it is framed as a problem, but not a problem for doctors, rather it is constructed as society’s problem.

B.7.5 Health is not a priority

health is and like a lot of people I would encounter who might have what I consider maybe not a great prognosis in life in terms of- P: their lifestyle P: there’s not always even loads of people aren’t willing to change but I think they don’t prioritise it P:whereas other people will I: will, yeah P: and I don’t really understand why that is
Doctor 1

Here we have the construction of health as something which can either be paid attention to or not, and the admission of ignorance from the respondent as to why this occurs.

you know, I think time people get sick because they’re not doing the basic things required to just keep your body in a modicum of health people just they don’t really care I suppose, or I wouldn’t even say care, they don’t want to get sick, but they downplay in their head the importance of kind of basic things like that, they really think that no, no, its fine if I don’t eat today Student 3

Here we have the same notion from a different perspective. Again, the lack of focus on health is constructed as the driver of this problem by both respondents. To a certain extent this reflects their pre-occupations, given that they are both focused on maintaining their own health, but there may be something behind these statements.

B.7.6 Linking Body and Mind

Again, this code only appears in the transcripts of alternative practitioners and students, and is conspicuous by its absence throughout the transcripts of the General Practitioners.

Some examples of this code follow below.

“ like I would press certain acupuncture points, work certain energy channels...

P: Through- which are directly applied on to the body so the focus of the person goes to those places – Alternative Therapist 1”

As one can see from this quote, this participant links physical manipulations of the body as having correlations to mental and emotional states. This is interesting as it is, in one sense, an extension of the idea of a healthy mind in a healthy body.

“Yeah. Well, in in in again with every organ in the body there’s there’s associated emotions and there’s pluses and minuses with that. So for example, we take an organ like the liver the liver ahhhhh If we’re totally overworked and ammm ehheh the liver can become very excitable and the energy of the liver it it it its a huge organ its it stores almost all of our blood – Alternative Therapist 3”

From this quote above, we can see a similar process to that observed from the first quote under this code heading. For this participant, emotional awareness can be correlated with the state of particular physical organs. Its interesting in that this is something which would probably never be done in psychology, as emotions would be regarded as being neural events rather than being distributed throughout the body.

“Same thing with the heart If the heart is healthy, you experience joy P:We find things funny, we like people, you know we think the best of every single person we meet P:When the heart is downtrodden and its there’s so many expressions – my heart was broken or my heart sank when I heard that news so (pause) even in the lairs of that – Alternative Therapist 3”

Again, we see a cognate description here, where emotions are embodied in the organs and orifices of the body, rather than dissociated from these in the brain/mind. Its perhaps a radical dissolution of the traditional dualism that pervades Western European thinking. For this respondent the locus of emotion is embodied in the very physical structures that make up the body, rather than being something that sits on

top of it, in the psychological or medical conceptions of emotion. It is interesting how this maps neatly to some of the theoretical conceptions expressed in Chapter 3, especially the embodied cognition section.

B.7.7 Linking Levels

This code is quite similar to the Linking Body and Mind code, but there are some important differences which we will explore below.

“trouble with knees, have often ehhhh, a [pause] a relationship to, to to, the what direction you take in life – Alternative Therapist 1”

“your hands P: what have you, how you handle things – Alternative Therapist 1”

“You know like your digestive system often is affected of how you digest life – Alternative Therapist”

We can see from these extracts that the respondent is linking problems of a physical nature to mental and emotional states. This excerpt tends to use argument by analogy which is not usually accepted as proof, but it does provide some insight into the construction of health by this participant. These quotes imply that the entire person is considered as one entity, and that any change in emotions, mental and physical issues will have an impact on all of the other areas.

“breathe our emotions – Alternative Therapist 3”

“Breathing emotions and leaving them out – Alternative Therapist 3”

We see here again evidence of an extremely holistic approach to the person, where emotional expression can be linked to breathing. It is perhaps interesting in this sense that an old word for life energy was prana, which was also the word for breath, A similar observation can be made about the Greek pneuma, which means both spirit and breath. The Hebrews also used the same word for breath as for emotions (Ruach), all of which points to the commonality of these ideas throughout history.

“probably equally important I presume it seems like kind of one feeds off the other if you’re physically sick, chances are you’re not going to be too well mentally, or if you’re mentally you’re probably not going to be too well physically
– Alternative Therapist 3”

Here we can see a recognition of the interdependence of the differing parts of the human organism and the idea that one of them can throw the others off their functioning.

they would also say our rivers are tributaries sorry our veins are like rivers, and capillaries like tributaries Alternative Therapist 3.

Here this theme focuses on the relationship between our body and the world, which tends to be a common theme in Chinese thought. There is a sense of reasoning from analogy, but the underlying notion is one of harmony with the environment and its continuity.

B.7.8 More Than Physical

ammmm I suppose you can look at it in different ways ammm I think it goes beyond biological Doctor 1

Here we have one of the respondents constructing health as not just the physical. This perhaps links to the conception of health as greater and more all-encompassing than sickness.

ammm a lot of people focus in I suppose the biological or physical issues

P: but usually they’re only a very small – well not a very small – but they’re only part of of the problem Doctor 1

Again here we see that others are constructed as focusing on the biological, while the respondent regards this as important but not the whole story. Note how the respondent first describes them as very small, and then goes back on this statement to a certain degree. Note again, the focus on “problems”.

and ammm you know certainly sometimes there comes a point with patients where you just have to say look there's no point giving you another drug- Doctor 1

- there's no biological reason for this so we need to explore alternatives P: alternative pathway Doctor 1

Here again, we have the framing as physiological explanations only going so far, but in this case it appears to be framed as an adjunct in the sense that only after the biological explanations are completed are other options considered.

yeah, but then you have a duty to to try to bring the patient to the understanding that because bio-medicine doesn't have the answer doesn't mean that your symptoms or your perspective is not valid and really Doctor 1

The respondent here ratchets back on the previous statement, making it clear that this should be framed in a patient-centred way, and not just in terms of hypochondria or something like that, but to respect the symptoms and try to get to the heart of the issue, wherever the causes may lie.

ahhh it it if used in itself it may not answer the whole problem because people come not only with a condition, they come with with a condition attached to their own bodies its got feelings, its got thoughts preconceptions, its got worries, etcetera- Doctor 2

The other doctors views this differently, or at least frames it in a different fashion - here the symptoms are the primary object of interest, but they are attached to people.

oh there's a lot more to it than that. Ammm obviously it is physical the physical aspects do lead to psychological and emotional ammm or both so there;s you know its definitely not just physical aspects

Student 2

Here again we have the focus on the physicality of health, but these are framed as impacting on other elements of health - note the use of the term "leads to", which seems to suggest this kind of reading.

B.7.9 Energy Model of Health

how do I think they work? Well....from doing different treatments....like shiatsu, so which works on the energetic plane Alternative Therapist 1

Here we have the matter-of-fact response that Shiatsu (a Japanese form of massage) works on the energetic plane. This plane is not particularly defined here, or indeed anywhere in the text.

you actually work energy channels that ammm...remind....the person to put energy in certain places. Yeah, like I would press certain acupuncture points, work certain energy channels...

This is a particularly interesting quote. In some sense, the body is seen as reflecting these energy channels, and the physical act of massage activates them. The most interesting word is remind, which suggests that this is a natural thing which is done, and not any effort on the part of the therapist.

it is too easy to say mind body soul but roughly its pretty much that, physical energy, taking the energy from your food and translating that into nutrition into energy and accommodating that and then expelling energy and expelling waste products efficiently that's the physical side of things energy then Alternative Therapist 3

Here we have the construction as one of different levels where the physical body is perceived as primary, taking energy (in presumably a physical sense) and nourishing the mind and soul. Note that this is again modelled as a cycle, food comes in and waste goes out.

physically taking in energy expending it, ammm clearing it out from the body and then ehhe emotional energy, and then mental energy same thing. - Alternative Therapist 3

Here the notion is more fully developed, that physical nourishment is analogous to mental and emotional nourishment, in that they follow the same cycle.

If one relation of all of those who take in the right energy your mental energy will be correct. There's emmm very common ahhe which my professor used to say – if

ahhh if you want to do exams, if you're studying, eat well. He says the spleen produces your blood your red blood cells

Here again, the same theme is developed in that the physical and mental are interdependent, a theme that also came through from the doctors, but expressed in a very different way. The mental is constructed as being completely inseparable from the physical, and indeed the organs of the body are imbued with certain elements "blood" in Chinese medicine tends not to mean the same thing as it does in Western paradigms.

In France they use digestifs, and apertifs, and that's that's the same thing, you know the same as preparing the energetics to assimilate....transform and bring forth the energy from that food very very carefully, and if you do that, then you're going to have Chi and if you've Chi the the character for Chi in Chinese medicine, or sorry in Chinese language ammm the I'll draw for you is is similar to that and basically what it symbolising is a fire, a pot with rice in the pot and steam rising from it so the pot is sitting over the fire cooking the soup

Here again, simple physical practices are constructed as being parallel to this system of energetics, and the use of allegory in this system is highlighted. The abstract notion of Chi is linked back to writing and also to an image, emphasising that this abstract notion is rooted in physical, common conceptions.

So again, when you say how do I think it works that's how I think it works as well manipulating energy P: ammm, the energy has to be there first day or first place, so someone is very depleted then often I have to make sure I that we get them to a place where they can that I can manipulate energy because if it's not there, I'm only going to deplete them much more

Here we can see the relationship between the two practitioners, in that the practitioner can only work with the patient, not actually imposing themselves upon them. The notion of energy is again referenced, but classified as a pre-existing thing within people, and that their energy is used to heal themselves. It is interesting to consider that this implies a model in which the therapist and patient are co-working on the treatment, which is a facet of treatment that is also expressed by the doctors, though in different language.

So, that's, in my thinking, in my study my training how I think things are working as well, that I am manipulating pockets of either static or depleted for

replete excessively replete some energy in the body and ahhh.... I: that effects change? P: Yeah, exactly, yeah, yeah. So that's I suppose my limited way of understanding it and the rest then is down to patient cooperation, really, go an

Again, this energy is described in very physical metaphors. The energy is stored in pockets, but the practitioner notes that this is merely a personal conception, but that the primary factor is patient co-operation.

, doctors, etcetera medications ammm because like a lot of the medication or the herbs there's there all like, natural, they're natural sources no chemicals or what have you so putting in a little bit that were, its just not, you know in ourselves and in our body ammmm and its just overall putting that Chi back into us also
Student 2

Here we have a similar conception of health, and especially this notion of energy. Again, the focus is put on self-healing and the notion of the patient as primary. There is a strong focus on "naturalness", this is constructed as being an important aspect of a treatment.

ammm its very very helpful for people that have say arthritis ammm could be a headache or ammm you know someone suffering from insomnia list goes on, but its just like I find it fascinating that it is is basically the universe is all around us, and that's basically how it works you know energy is coming from the universe and its flowing in through ammm the practitioners hands the recipient without having not actually having to touch Student No 2

Here again, we have a similar conception - the universe is all around us, the power does not come from within ourselves, but rather healing is a property that can be channeled by one person towards another, but in fact is not caused by either of them.

but ammm I I personally find that its its you know it does alleviate pain you know to the highest degree from my own experience ammm and then reflexology is an old Chinese medicine and its – well not medicine but its- I:diagnosis
P:-yeah, exactly and you have if you can imagine your feet have meridians travelling from your feet that connect with each part of your body and I find that in fact to be my favourite alternative therapy

The respondent was here talking about acupuncture. Again, the model of energy is raised, only after talking about how it produces real changes for the respondent. It

again is described as this abstract energy that is legitimised by its effects on the physical.

B.8 Personal Factors

The next major section is on personal factors in health.

B.8.1 Awareness

If a person is not willing to work on themselves and to get better, you can do whatever you want, and they- I: they won't P: and I: so it requires their willingness to take part in the procedure as well- P: yes, their awareness-
Alternative Therapist 1

Here again, we can see a similar theme as above, where the person undergoing treatment is perceived as primary, and the therapist as an adjunct to this.

If your hemoglobin your iron is low, you don't absorb information as good just to – they say that the thing if you errr if you're studying and you're eating while you're even studying trying to combine the two at a time (pause) there's an expression in Chinese as well to say that you might as well pick up the book and eat it, cos you'll get more nutrition from eating the paper then you will from eating your meal if you don't have conscious intent Alternative Therapist 3

Here we have the notion that awareness is the key to gaining nourishment from the physical world. Again, this is constructed as wisdom from others, but there is also a sense of keeping these levels distinct, and focusing on what it is you are doing in the present moment.

By paying attention to what is (pause) going on with yourself.

Alternative Therapist 3

actually yeah, no the entire thing making people aware of the fact of how – how skewed our perceptions of health are a bit cos when I kind of realised it it made me feel about my health quite a bit differently Student 3

B.8.2 Construction of Illness

but usually they're only a very small – well not a very small – but they're only part of of the problem a lot of it is how people will approach that physical problem- Doctor 1

Here we can see the manner in which the doctors construct this notion of needing the buy-in from the patient. It is framed as if the way in which someone approaches the problem is key to the outcome.

-if they come with a physical issue- P:- it would be how they actually understand it, perceive it, the effects it's having on their life ammm the social supports they have around them to deal with it ammm and the effect it will have on them socially in terms of their ability maybe to manage at home to work- Doctor 1

Here again, the patients perception of their problem and its treatment is regarded as important, as is their connections to the social environment.

that sort of thing, so you know ammm yeah, so I:Its kinda how they interpret it forms a large part of it P:yeah, its very much how the patient interprets it- P:the problem Doctor 1

Here again we have a similar kind of statement as coming from the alternative therapists, though constructed with a different lens, the lens of social context. Nonetheless, again, a large burden is placed on the patient by both groups, though in a different fashion.

I think your perception of sickness is so different, a lot – sometimes people come to me with what I consider total rubbish and that obviously is quite – you don't say that to them obviously – I acknowledge their concerns, ideas and expectations but you know really in the grand scheme of life its like get on with it- P: - whereas when you actually see people who are sick- P: - it just puts health in perspective maybe Doctor 1

Here we have how the respondent constructs health themselves, as a result of their experiences with patients. Note that this is framed as making the respondent less likely to worry about their own health, and also their experience of actual sickness makes them less likely to see some reports as valid problems.

B.8.3 Different Needs for Different People

I think they're good, I think we all work with different groups Alternative
Therapist 2

Here we have one of the therapists talking about doctors, note that this is framed in terms of personal choice, in that different forms of health care are modeled as being for different groups of people.

Full of different amounts of ammm you know the common misconception of you need everybody needs eight glasses of water a day Maybe I do, maybe you don't If if if someone puts too much water into their garden the roots sort of become loose and they'll come out too easily. So, in in a some people they need far more water than that. So its such its such a personal thing per each individual working out what what needs addressing in their garden, basically - Alternative Therapist
3

Here we have a similar theme, though expressed in a different fashion. Here the focus is constructed as there being different treatment needs, rather than therapist needs. Note that this is again framed (as is common from this respondent) with environmental/natural metaphors.

I always like to be able to gatekeep for people as to what is the most appropriate treatment for them and I think that's a really really good thing Alternative
Therapist 3

This particular excerpt is interesting in that it places the respondent in a position of power above the patient, by "gatekeep"-ing their particular forms of treatment. This is interesting in that it is a theme which comes through in the responses of doctors, but not so much in the responses of therapists.

you know that sort of relationship with somebody and you know there's some people I think its like this sort of phrase horses for courses there are some patients who'll get on with certain types of doctors and vice versa Doctor 1

Again we have a construction of different needs for patients which is subtly different from the ones above. Here the construction is that the relationship between patient and practitioner is of primary importance.

ammmm because one person might go down the road of therapy but because there may be a chemical imbalance then they may need ammm whether it be alternative medication or ammm you know anti-depressants or something like that ammm sometimes I feel if one doesn't work you should definitely need the other Student 2

The final excerpt from this section focuses on the use of different forms of treatment as differentially impactful on different people. Note that health care is framed in a somewhat "a-la-carte" way, in that if one form of treatment doesn't work, then the person should definitely attempt the other.

I haven't heard any particularly negative things people get a lot of good work out of it get a lot of good out of it, things like homeopathy which has no real scientific backing whatsoever well from my viewpoint ammm...seems to have quite a lot of people who feel a lot better afterward Student 2

B.8.4 Expectations Differ Across Class

I'm not saying that everything is necessarily is good either, but the barriers that are automatically there are whereas if say to somebody maybe from a higher socio-economic group you know really your weight is the biggest problem- P: for you you know you need to work on this- P:- and look at your lifestyle, your day and how we can work this out ammm and we might talk about gym or personal trainer and they'll be much more....you know....they just sort of yeah, yeah, my friends do that, or I could do this they don't sort of come with a negative- I:yeah P:-viewpoint sometimes you encounter from other people. And I don't think the other people are being negative, I just think that there are practical realities of their life Doctor 1

Here again we have the focus on class, where the social and environmental situation surrounding people makes it easier or harder to focus on fixing their problems. Note again that typically of the doctors, the focus is placed on the problem, rather than the more holistic viewpoint (which, to be fair, is expressed by the doctors in other excerpts).

B.9 Models of Sickness

B.9.1 Sickness Part of Life

I don't really do much for it the most standard thing is if I get sick I'll make an effort to eat better more regularly, take more fluids, lie down more its all very basic things – in I wouldn't really go for a medical treatment unless I'd be really there's something particularly like as an emergency mostly out of the fact that I tend to view these things as minorly tieing into the thing of you know ill health – I don't really view unhealth as particularly aberrant I suppose. Yeah, no I standard thing would just be to take bed rest take a very very standard boring but am if seriously ill I'd go to a GP or doctor. Student 2

Here is an excerpt from one of the students who were interviewed. Note that sickness is contextualised as quite normal, regarding them as part of typical life. The respondent says that they would focus on bed rest, but if “seriously ill” would go see a doctor. This notion of seriousness crops up in the interviews, but it is not clearly defined at any point.

I tend to very much go on a go about my daily day the way I would normally -
P: - because that tends to make me feel a lot better actually kinda kinda go oh I'm sick and sit down for the day and try to recover a lot of the time I actually feel a lot more sick than in a similar circumstance where I had to go out and interact with the world I just by going about my daily day like that keeping to my routine, by basically ignoring how ill I can be - Student 2

Here again we have the same theme repeated, in that sickness is not framed as debilitating, rather it is the actions of “being sick” that are framed as being the real problem. Note that this respondent was quite young and healthy, and so has probably not experienced any major health emergencies which would not respond well to this form of treatment.

B.9.2 Sickness from Pathogens

I mean obvious ones of course are going to be things like you know, pathogens of some descriptions, like like whatever they are, bacteria or viruses etcetera apart there there's an interaction there of course the integrity of the person's immune system- Doctor 2

Here is the standard bio-medical model of illness, whereby sickness is conceptualised as coming from pathogens and their interaction with the body's immune system.

physically viruses, bacteria what in....in very, well I suppose like is the obvious what actually sickness is where sickness comes from Student 3.

Here is the same model, expressed by one of the students. Note that it is "obvious" to this respondent, though all of the doctors and practitioners are much more measured around this subject.

B.9.3 Sickness as a Way of Life

in a manner that gonna help you, you know and then there are a group of patients who don't actually really want cure and and that's the other sort of side of the coin- P:- that you're not going certainly not anymore you're not going for cure in a lot of cases ammm and then there's some people who are sort of eager to stay in a certain perspective- Doctor 2

Here we have sickness conceptualised as something which people actually fight to maintain — this group of people who don't really want a cure. The other framing in this section is that for many of the illnesses seen by the doctor, there are no good cures, merely maintenance treatments.

B.9.4 Services focus on Illness

I suppose we're more set up in our training to deal with ill-health as a concept- I:yeah P:- than health and particularly how the health services are structured that focus on- I:mmmm P:- ill-health rather than health promotion. Which is a very very different type of engagement with with the person, I think Doctor 1

Here we see that the health services are framed as being focused on illness, as is the entire training and socialisation of doctors.

and its because we're stuck in the ill-health- I:mmmm P: setup and not in the health set-up

Note here that it is framed as a large problem "Stuck" in the ill-health setup, and not the health setup. Note that this is framed as being a systemic problem, which is a viewpoint which has some merit, but also serves to reduce the responsibility of the practitioners. It is perhaps somewhat determined by the training which doctors tend to receive.

B.9.5 Ill-health as Normal

they do a lot of damaging activities kind of are so used to it, and consider it such an integral – or such a normal part of their life that they don't really view the ill-health they get as a consequence- I:mmmm P: - in any way abnormal Student 3

Here we have ill-health being framed as the default state of affairs for many people. It is constructed as both stemming from a lack of awareness of consequences, and as the result of not knowing what health is like. This again reinforces the theme of health/sickness as being personally mediated, which comes through in all respondents, regardless of background.

bit. People – people have ill health very much, rather like ammm people aren't particularly healthy I mean even in – even in my earlier it was just an absence of particular ill health, I mean I don't consider I consider being in a relative state of fitness and not particularly healthy to be the normal rather than actually being healthy Student 3

Again, here sickness is framed as both the absence of health, and also as a lack of awareness of what it means to be healthy. Sickness is framed by this respondent as being the natural state of affairs, rather than something unnatural. In fact, health is framed somewhat as the more unlikely feature of life.

B.9.6 Differences in Immunity

and so basically those with any type of weakened immunity will be more susceptible if those pathogens get into the body because a lot of us go round with pathogens which wouldn't cause very serious illness in and of themselves but the immune system is dealing with them, and we've got effective barriers to that that infection. So I suppose that there's there's a number of issues and I suppose that the immunity and sometimes that immunity can be modulated by by by a person's mental well-being, the levels of stress can weaken that Doctor 2

Here we have the way in which the mind is perceived to affect the body framed by one of the doctors. Note that this is all very qualified, the use of qualifiers such as "I suppose", and "sometimes" makes it clear that this is not a particularly strong viewpoint from this respondents perspective. Again, the body is regarded as primary, with the mental and social worlds of the patient regarded as an afterthought.

B.10 Social Dimensions

B.10.1 Class Differences

say if I've a 45 year old man who comes into me with chest pain and I think its a cough, and he's been smoking and he is working as maybe a builder and he's 4 kids at home and.... you know, that's sort of where he's at....to promote his health takes huge efforts because he has to invest much more in making changes to his diet, his smoking habits and his general social milieu, if you like Doctor 1.

This excerpt frames health and health problems and the ability for them to be resolved as inextricably linked to the social aspects of this life. The social environment around the person is regarded as something which can have a causal impact on outcomes.

whereas if I have the same man coming from maybe a higher socio-economic group he's expectations of his health are automatically higher- I:mmmm P: and he's probably gonna have the leisure time or the money to access the appropriate resources- I:yeah P: that are just not available to say my poorer-

Here we have the opposite social situation, but with similar symptoms. It is interesting to note that the "expectations" around health are regarded as better, given

that this construct is used throughout placebo research. Again, the focus is on the nature and quality of the social resources available to the patient rather than any physical measure which makes it more likely that one rather than the other would succeed.

So for example they did this sort of exercise referral programme here running in cork- I:right, OK P: - So I can say to somebody, look I'm gonna refer to the gym and ammm they'll assess your fitness and then you'll get a reduced rate to the gym- I:mmmm P: - and that sounds good but my patients will come back and say ammm I don't have the bus fare to get to the gym of course I would say then you could walk you know I can't transport to the gym is difficult and a lot of the people who we would be referring might have a lot of health issues you know like maybe a bit of arthritis, chest problems- I:mmmm P:- so me saying to them walk is just not a runner especially in the rain and you need gear and all that sort of stuff. The second thing then is the waiting list the third problem is is reduced rate, but its not free-

ammm, I'm not saying that everything is necessarily is good either, but the barriers that are automatically there are whereas if say to somebody maybe from a higher socio-economic group you know really your weight is the biggest problem- I:mmmm P: for you you know you need to work on this- I:mmmm P:- and look at your lifestyle, your day and how we can work this out ammm and we might talk about gym or personal trainer and they'll be much more.....you know....they just sort of yeah, yeah, my friends do that, or I could do this they don't sort of come with a negative- I:yeah P:-viewpoint sometimes you encounter from other people. And I don't think the other people are being negative, I just think that there are practical realities of their life Doctor 1

Again here, the social context of the medical advice is regarded as a primary driver of the outcomes, in that the poorer patients have both lower expectations and less resources to help them to achieve the health goals. The phrase "practical realities" both sums up the injustice and references the idea introduced earlier that people have different capacity to actually maintain their health, and most of these factors are entirely outside the control of a health practitioner.

:yeah but it it its more than time people from higher socio-economic groups can be working very hard- I:mmmm P: as well like but they value health...they can see the value of health more I:more, yeah P:- I think you know I:yeah P:ammm, so....time is an issue, cost is an issue but its also perception of how important- I:yeah P:- health is Doctor 1

Again here, the issue is framed as not just one of resources, but rather that people from higher socio-economic groups perceive their health as more important or more pressing than do other people.

you know and I can nearly you can sadly nearly predict I: who's going to P: or
yeah you can predict ill health from a young age or you can predict
sometimes.....people that are more likely to engage encounter maybe trouble

Here the issue is framed as a sad fact around the work which the respondent does with people, in that it is easy to see based on the particular socio-economic status whether or not treatment will be effective. Again, there is a subtext of framing this as a problem for society rather than just for doctors.

And it comes right down to issues such as you know, equity and and even the imbalance of wealth as well, that's not an easy one to solve well there's a lot of ill health and disease is is very much linked with poverty and you can see the differences across socio-economic groups Doctor 2

Here we have a similar framing of the problem by the other doctor respondent, again there is somewhat more of a framing of this as not something which can be changed, rather it is an unfortunate fact of life for this respondent.

yes, it can definitely or it can be just effected through the sufficient funds to go for the cheapest option all the time, and that might not be the healthiest I:the healthiest option P: and in terms of choice of foods, for example, maybe going for the take-away or maybe going for, or the skills may not be there to actually produce cook the proper food. Doctor 2

Here it is framed both in terms of money — “the cheapest option” and also in terms of a lack of resources to actually prepare the food. Again, it is framed in a particularly abstract fashion, note the use of the passive voice, which is a common pattern for this respondent.

I don't really know why that is its not only money I:yeah P:It is – it is a mindset
I think too Doctor 1

Again, here the problem is framed not only in terms of access to monetary resources, but also as a “mindset”, which is something which falls outside the purview of doctors.

This framing fits quite well with earlier commentary about health as a social problem.

B.10.2 Family dimension of illness

because a lot of the problems I deal with have been ammm people are nearly reared to them you know. Doctor 1

Here again, we have many of the problems encountered by the doctor in their practice framed as being more social than medical. There is a sense in which the environment in which people live is regarded as a major contributing factor towards health or the lack thereof.

It is Its a silly way like the health seeking behaviour I have in some of my patients is just totally reared into them I can see it being reared into them I:mmmm P: and some I spend more time on each visit trying to convince them people not to use the antibiotic for example ammm....but they pick up the anxieties from of their family

Doctor 1

This excerpt frames the issue as something which is acquired from the social environment, whereby patients demand antibiotics because their family say they need them. There is a certain way in which this is framed as a reason for differential outcomes amongst similar patients, and also as a device by which to explain away poor outcomes.

its very frustrating working in a system and I'm that's reared into people the whole depression issue meds are not you know in common parlance the only way to deal with things and then counselling is just too nebulous Doctor 1

Here we can sense the frustration (explicitly called out in the text) that patients come with their own expectations around treatment, and in some cases the respondent just has to go along with this, to gain any hope of a successful outcome.

the origins of ill health, as it were come right from the household level Doctor 2.

Here we can see a similar point being raised by another doctor respondent. It is perhaps interesting to note that family systems are regarded as primary in many cases, the use of the term “reared” into them. It is as though the doctors allow themselves to consider some social issues, but only to the family level. Alternatively, this particular framing could be the result of their training only considering the systems up to this level and not wanting to speculate beyond the data.

So for example they did this sort of exercise referral programme here running in cork- I:right, OK P: - So I can say to somebody, look I'm gonna refer to the gym and ammm they'll assess your fitness and then you'll get a reduced rate to the gym- I:mmmm P: - and that sounds good but my patients will come back and say ammm I don't have the bus fare to get to the gym of course I would say then you could walk you know I can't transport to the gym is difficult and a lot of the people who we would be referring might have a lot of health issues you know like maybe a bit of arthritis, chest problems- I:mmmm P:- so me saying to them walk is just not a runner especially in the rain and you need gear and all that sort of stuff. The second thing then is the waiting list the third problem is is reduced rate, but its not free- Doctor 1

ammm, and I suppose if we look at each such as a a part of the effect I suppose, you know, we can also say that link stress to some degree with habits and dangerous habits so to speak such as smoking. Its a two way chan- thing and one can give rise to the other. Twill often be ahh if we look at it, I mean poverty too can give rise to ill health

B.11 Conclusions

From this part of the chapter, it can be seen that doctors and alternative therapists had very different ways of conceptualising health. The alternative therapists tended to regard themselves as facilitators, while the doctors appeared to regard themselves (or at least be regarded as this by others) as having more control and authority than the patient. The transcripts of the students tended to reinforce rather than change these themes.

It can also be seen that alternative therapists tended to conceive as health being “natural” and being in tune with one’s body, while doctors tended to focus on the physical body, and also appeared to have more of a focus on the problems of society than did the alternative therapists. This makes sense, given that doctors are more

likely to encounter a wider variety of patients, due to the cost associated with alternative treatments.

One issue that came through from all the respondents is that health was conceptualised as being much broader than was sickness. Sickness was regarded as a problem, while health appears to require more effort, and that this effort could not be down to the individual alone.

With regards to the development of the IAT, this research focused the mind of the researcher. Coming through in many of the themes was the different outlook between doctors and alternative treatments. They appeared to reflect two quite distinct ways of approaching treatment and health more generally. Therefore, one way in which to capture this variance in the form of an implicit measure was to make the contrast between the two forms of treatment the salient dimension on which participants would classify. This subject is explored in much more detail in Chapter 6.

Appendix C

Appendix III: Health Rep Grids

C.1 Health Rep Grids and Sorts

C.1.1 Instructions

1. Using the sheet provided, write in the name of a person who fits the category described. The researcher is not interested in this page, and you may take it with you when you are finished
2. If you cannot think of a person who completely fits the category, use someone who is as close as possible to the role.
3. Do not re-use people for different roles.
4. Secondly, look at the other sheet. Note the roles along the right hand side of the page
5. In the gender column, write M if the person you associate with this role is male, and F if this person is female.
6. Thirdly, look at the column labeled sort. This tells you which three roles to combine
7. The first sort says 1, 2, 3. These numbers reference the roles. In this case, you are asked to compare a health professional whom you liked, an exercising person and a health professional whom you disliked
8. In the similarity column, write how two of these figures are similar. In the difference column , write how one of them is different.
9. In the other columns, put an X if the similarity construct applies, and an O if the difference construct applies. Write nothing if the constructs are not relevant
10. Continue down the page until all sorts have been completed.

Table C.1.1 shows the sorts that participants were asked to carry out. Table C.1 shows the form in which participants filled out the respective people from their lives (this form was never shared with the researcher).

	Role	Gender	Sort	Similarity	Difference	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14
1	Health Professional whom you liked		1,2,3																
2	An exercising person		2,3,4																
3	Health Professional whom you disliked		3,4,5																
4	A partner who was/is unhealthy		4,5,6																
5	Alternative Practitioner		5,6,7																
6	Yourself		6,7,8																
7	A healthy family member		7,8,9																
8	An unhealthy family member		8,9,10																
9	Someone who cares about your health		9,10,11																
10	An unhealthy friend		10,11,12																
11	Someone who gives good health advice		11,12,13																
12	A happy person		12,13,14																
13	A healthy friend		13,14,1																
14	A partner who was/is healthy		14,1,2																

Table C.1: List of People

	Role	Name
1	Health Professional whom you liked	
2	An exercising person	
3	Health Professional whom you disliked	
4	An unhealthy former or current partner	
5	Alternative Health Practitioner	
6	Yourself	
7	A healthy family member	
8	An unhealthy family member	
9	A person who cares about your health	
10	An unhealthy friend	
11	Someone who gives good advice	
12	A happy person	
13	A healthy friend	
14	A healthy former or current partner	
15	An unhappy person	

References

- Ackerman, J. M., Nocera, C. C., & Bargh, J. A. (2010). Incidental haptic sensations influence social judgments and decisions. *Science*, 328(5986), 1712–1715.
- Ahern, A. L., Bennett, K. M., & Hetherington, M. M. (2008). Internalization of the ultra-thin ideal: Positive implicit associations with underweight fashion models are associated with drive for thinness in young women. *Eating Disorders: The Journal of Treatment & Prevention*, 16(4), 294-307.
- Akaike, H. (1974). A new look at the statistical model identification. *Automatic Control, IEEE Transactions on*, 19(6), 716–723.
- Allan, L. G., & Siegel, S. (2002). A signal detection theory analysis of the placebo effect. *Eval Health Prof*, 25(4), 410-420.
- Allison, P. J., Guichard, C., & Gilain, L. (2000). A prospective investigation of dispositional optimism as a predictor of health-related quality of life in head and neck cancer patients. *Quality of Life Research*, 9(8), 10p.
- Amanzio, M., & Benedetti, F. (1999). Neuropharmacological dissection of placebo analgesia: Expectation-activated opioid systems versus conditioning-activated specific subsystems. *Journal of Neuroscience*, 19(1), 484-494.
- Amanzio, M., Pollo, A., Maggi, G., & Benedetti, F. (2001). Response variability to analgesics: a role for non-specific activation of endogenous opioids. *Pain*, 90(3), 205-215.
- Andersson, G. (1996). The benefits of optimism: A meta-analytic review of the Life Orientation Test. *Personality and Individual Differences*, 21(5), 719–725.
- Arcuri, L., Castelli, L., Galdi, S., Zogmaister, C., & Amadori, A. (2008). Predicting the vote: Implicit attitudes as predictors of the future behavior of decided and undecided voters. *Political Psychology*, 29(3), 369-387.
- Asendorpf, J. B., Banse, R., & Miercke, D. (2002). Double dissociation between implicit and explicit personality self-concept: The case of shy behavior. *Journal of Personality and Social Psychology*, 83(2), 380-393.
- Baker, S. R. (2007). Dispositional optimism and health status, symptoms and behaviours: Assessing idiographic relationships using a prospective daily diary approach. *Psychology & Health*, 22(4), 25p.

- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, 84(2), 191-215.
- Barrett, P. (2007). Structural equation modelling: Adjudging model fit. *Personality and Individual Differences*, 42(5), 815-824.
- Bausell, R. B., Lixing, L., Bergman, S., Wen-Lin, L., & Berman, B. M. (2005). Is acupuncture analgesia an expectancy effect? *Evaluation & the Health Professions*, 28(1), 9-26.
- Beecher, H. (1955). The powerful placebo. *Jama*, 159(17), 1602.
- Benedetti, F. (1996). The opposite effects of the opiate antagonist naloxone and the cholecystikinin antagonist proglumide on placebo analgesia. *Pain*, 64(3), 535-543.
- Benedetti, F. (2005). The importance of considering the effects of perceived group assignment in placebo-controlled trials. *Evaluation & the health professions*, 28(1), 5-6.
- Benedetti, F. (2007). What do you expect from this treatment? changing our mind about clinical trials. *Pain*, 128(3), 193-194.
- Benedetti, F. (2008). Mechanisms of placebo and placebo-related effects across diseases and treatments. *Annual Review of Pharmacology & Toxicology*, 48(1), 33-60.
- Benedetti, F., Amanzio, M., Baldi, S., Casadio, C., & Maggi, G. (1999). Inducing placebo respiratory depressant responses in humans via opioid receptors. *European Journal of Neuroscience*, 11(2), 625.
- Benedetti, F., Amanzio, M., & Maggi, G. (1995). Potentiation of placebo analgesia by proglumide. *The Lancet*, 346(8984), 1231-1231.
- Benedetti, F., Amanzio, M., Vighetti, S., & Asteggiano, G. (2006). The biochemical and neuroendocrine bases of the hyperalgesic nocebo effect. *Journal of Neuroscience*, 26(46), 18-18.
- Benedetti, F., Arduino, C., & Amanzio, M. (1999). Somatotopic activation of opioid systems by target-directed expectations of analgesia. *Journal of Neuroscience*, 19(9), 3639-3648.
- Benedetti, F., Mayberg, H. S., Wager, T. D., Stohler, C. S., & Zubieta, J.-K. (2005). Neurobiological mechanisms of the placebo effect. *Journal of Neuroscience*, 25(45), 10390-10402.
- Benedetti, F., Pollo, A., Lopiano, L., Lanotte, M., Vighetti, S., & Rainero, I. (2003). Conscious expectation and unconscious conditioning in analgesic, motor, and hormonal placebo/nocebo responses. *Journal of Neuroscience*, 23(10), 4315-4323.
- Benedetti, F., Rainero, I., & Pollo, A. (2003). New insights into placebo analgesia. *Current opinion in anaesthesiology*, 16(5), 515 (6 pages).
- Bentler, P. (1990). Comparative fit indexes in structural models. *Psychological bulletin*, 107(2), 238.

- Blanton, H., & Jaccard, J. (2006). Arbitrary metrics in psychology. *American Psychologist*, 61(1), 27-41.
- Blanton, H., Jaccard, J., Gonzales, P. M., & Christie, C. (2006). Decoding the implicit association test: Implications for criterion prediction. *Journal of Experimental Social Psychology*, 42(2), 192-212.
- Boker, S., Neale, M., Maes, H., Wilde, M., Spiegel, M., Brick, T., . . . Fox, J. (in press). Openmx: An open source extended structural equation modeling framework. *Psychometrika*. (In press)
- Boldero, J. M., Rawlings, D., & Haslam, N. (2007). Convergence between gnat-assessed implicit and explicit personality. *European Journal of Personality*, 21(3), 341-358.
- Bootzin, R. R. (2003). Studying the context in which treatments are delivered: Observations on 'open versus hidden medical treatments'. *Prevention & Treatment*, 6(1), 3c.
- Boysen, G. A., Vogel, D. L., & Madon, S. (2006). A public versus private administration of the implicit association test. *European Journal of Social Psychology*, 36(6), 845-856.
- Branthwaite, A., & Cooper, P. (1981). Analgesic effects of branding in treatment of headaches. *British Medical Journal*, 282(6276), 1576-1578.
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. In (Vol. 3, pp. 77-101). Taylor & Francis.
- Breiman, L. (2001). Statistical modeling: The two cultures. *Statistical Science*, 199-215.
- Brown, & Ryan, R. (2003). The benefits of being present: Mindfulness and its role in psychological well-being. *Journal of personality and social psychology*, 84(4), 822-848.
- Brown, Ryan, R., & Creswell, J. (2007). Addressing fundamental questions about mindfulness. *Psychological Inquiry*, 18(4), 272-281.
- Brown, C., Battista, D. R., Bruehlman, R., Sereika, S. S., Thase, M. E., & Dunbar-Jacob, J. (2005). Beliefs about antidepressant medications in primary care patients: relationship to self-reported adherence. *Medical care*, 43(12), 1203-1207.
- Carlson, L. E., Speca, M., Faris, P., & Patel, K. D. (2007). One year pre-post intervention follow-up of psychological, immune, endocrine and blood pressure outcomes of mindfulness-based stress reduction (mbsr) in breast and prostate cancer outpatients. *Brain, Behavior & Immunity*, 21(8), 12p.
- Carmody, J., Reed, G., Kristeller, J., & Merriam, P. (2008). Mindfulness, spirituality, and health-related symptoms. *Journal of Psychosomatic Research*, 64(4), 11p.
- Carson, J. W., Carson, K. M., Gil, K. M., & Baucom, D. H. (2004). Mindfulness-based relationship enhancement. *Behavior Therapy*, 35(3), 24p.
- Carver, C. S., Scheier, M. F., & Segerstrom, S. C. (2010). Optimism. *Clinical*

- Psychology Review, In Press, Corrected Proof.*
- Caspi, O., & Bootzin, R. R. (2002). Evaluating how placebos produce change. *Evaluation & the Health Professions, 25*(4), 436.
- Castelli, L., & Tomelleri, S. (2008). Contextual effects on prejudiced attitudes: When the presence of others leads to more egalitarian responses. *Journal of Experimental Social Psychology, 44*(3), 679-686.
- Chiesa, A., & Serretti, A. (2009). Mindfulness-based stress reduction for stress management in healthy people: A review and meta-analysis. *Journal of Alternative & Complementary Medicine, 15*(5), 8p.
- Cohen, J. (1988). Statistical power analysis for the behavioral sciences.
- Colloca, L., & Benedetti, F. (2006). How prior experience shapes placebo analgesia. *Pain, 124*(1-2), 126-133.
- Colloca, L., Sigauco, M., & Benedetti, F. (2008). The role of learning in nocebo and placebo effects. *Acute Pain, 10*(2), 102-102.
- Concato, J., Shah, N., & Horwitz, R. I. (2000). Randomized, controlled trials, observational studies, and the hierarchy of research designs. *New England Journal of Medicine, 342*(25), 1887-1892.
- Conner, & Barrett, L. F. (2005). Implicit self-attitudes predict spontaneous affect in daily life. *Emotion, 5*(4), 476-488.
- Conner, M. T., Perugini, M., O’Gorman, R., Ayres, K., & Prestwich, A. (2007). Relations between implicit and explicit measures of attitudes and measures of behavior: Evidence of moderation by individual difference variables. *Personality and Social Psychology Bulletin, 33*(12), 1727-1740.
- Conway, F., Magai, C., Springer, C., & Jones, S. C. (2008). Optimism and pessimism as predictors of physical and psychological health among grandmothers raising their grandchildren. *Journal of Research in Personality, 42*(5), 1352-1357.
- Craeynest, M., Crombez, G., Koster, E. H. W., Haerens, L., & De Bourdeaudhuij, I. (2008). Cognitive-motivational determinants of fat food consumption in overweight and obese youngsters: The implicit association between fat food and arousal. *Journal of Behavior Therapy and Experimental Psychiatry, 39*(3), 354-368.
- Crow, R., Gage, H., Hampson, S., Hart, J., Kimber, A., & Thomas, H. (1999). The role of expectancies in the placebo effect and their use in the delivery of health care: a systematic review. *Health Technol Assess, 3*(3), 1-96.
- Crum, A. J., & Langer, E. J. (2007). Mind-set matters: Exercise and the placebo effect. *Psychological Science, 18*(2), 165-171.
- Cwir, D., Carr, P. B., Walton, G. M., & Spencer, S. J. (2011). Your heart makes my heart move: Cues of social connectedness cause shared emotions and physiological states among strangers. *Journal of Experimental Social Psychology, 47*(3), 661-664.
- Dasgupta, N., & Greenwald, A. G. (2001). On the malleability of automatic

- attitudes: Combating automatic prejudice with images of admired and disliked individuals. *Journal of Personality and Social Psychology*, 81(5), 800-814.
- Dawes, R. (1979). The robust beauty of improper linear models in decision making. *American psychologist*, 34(7), 571.
- De Houwer, J., Beckers, T., & Moors, A. (2007). Novel attitudes can be faked on the implicit association test. *Journal of Experimental Social Psychology*, 43(6), 972-978.
- De Pascalis, V., Chiaradia, C., & Carotenuto, E. (2002). The contribution of suggestibility and expectation to placebo analgesia phenomenon in an experimental setting. *Pain*, 96(3), 393-402.
- Devilley, & Borkovec, T. (2000). Psychometric properties of the credibility/expectancy questionnaire. *Journal of Behavior Therapy and Experimental Psychiatry*, 31(2), 73-86.
- Devilley, G. J. (2002). The psychological effects of a lifestyle management course on war veterans and their spouses. *Journal of Clinical Psychology*, 58(9), 1119.
- Devilley, G. J., & Huth, A. (2008). Perceived distress and endorsement for cognitive- or exposure-based treatments following trauma. *Australian Psychologist*, 43(1), 7-14.
- DiBlasi, Harkness, E., Ernst, E., Georgiou, A., & Kleijnen, J. (2001). Influence of context effects on health outcomes: a systematic review. *The Lancet*, 357(9258), 757-762.
- DiBlasi, Kaptchuk, T. J., Weinman, J., & Kleijnen, J. (2002). Informing participants of allocation to placebo at trial closure: postal survey. *British Medical Journal*, 325(7376), 1329.
- Dictionary, O. E. (2010). (1989). Retrieved from OED Online. Accessed August, 30.
- Egloff, B., & Schmukle, S. C. (2003). Does social desirability moderate the relationship between implicit and explicit anxiety measures? *Personality and Individual Differences*, 35(7), 1697-1706.
- Egloff, B., Schwerdtfeger, A., & Schmukle, S. C. (2005). Temporal stability of the implicit association test-anxiety. *Journal of Personality Assessment*, 84(1), 82-88.
- Ernst, E., & Resch, K. L. (1995). Concept of true and perceived placebo effects. *BMJ: British Medical Journal*, 311(7004), 551-553.
- Evans, D. (2002). Pain, evolution, and the placebo response. *Behavioral and Brain Sciences* 25, no. 4, 459-460 (2 pages).
- Evans, D. (2003). *Placebo: The belief effect*. HarperCollins.
- Fabrigar, L., Wegener, D., MacCallum, R., & Strahan, E. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological methods*, 4, 272-299.
- Fernandez, E., & Turk, D. C. (1994). Demand characteristics underlying differential ratings of sensory versus affective components of pain. *J Behav Med*, 17(4),

375-90.

- Ferreira, P. L. (2000). [development of the portuguese version of mos sf-36. part i. cultural and linguistic adaptation]. *Acta Med Port*, 13(1-2), 55-66.
- Finlay, L. (2002). Outing the researcher: The provenance, process, and practice of reflexivity. *Qualitative Health Research*, 12(4), 531-545.
- Finniss, D. G., & Benedetti, F. (2005). : The neural matrix of pain processing and placebo analgesia: Implications for clinical practice. *Headache Currents* 2, no, 6, 132-138.
- Fischer, G., & Molenaar, I. (1995). *Rasch models: Foundations, recent developments, and applications*. Springer.
- Flaten, M. A., Simonsen, T., & Olsen, H. (1999). Drug-related information generates placebo and nocebo responses that modify the drug response. *Psychosom Med*, 61(2), 250-5.
- Flaten, M. A., Simonsen, T., Zahlse, K., Aamo, T., Sager, G., & Olsen, H. (2004). Stimulant and relaxant drugs combined with stimulant and relaxant information: a study of active placebo. *Psychopharmacology*, 176(3-4), 426-34.
- Friedman, J., Tibshirani, R., & Hastie, T. (2009). *The elements of statistical learning: Data mining, inference, and prediction*. Springer-Verlag New York.
- Fujita, T., Fujii, Y., Nakamura, T., Miyauchi, A., & Takagi, Y. (2000). Effect of avicatonin (chicken carbocalcitonin) on galvanic skin response: a randomized, prospective, double-blind, controlled study for an objective assessment of pain. *Calcif Tissue Int*, 66(4), 243-7.
- Gaitan-Sierra, C., & Hyland, M. E. (2011). Nonspecific mechanisms that enhance well-being in health-promoting behaviors. *Health Psychology*, 30(6), 793.
- Gawronski, B., & Bodenhausen, G. V. (2005). Accessibility effects on implicit social cognition: The role of knowledge activation and retrieval experiences. *Journal of Personality and Social Psychology*, 89(5), 672-685.
- Geers, A. L., Helfer, S. G., Kosbab, K., Weiland, P. E., & Landry, S. J. (2005). Reconsidering the role of personality in placebo effects: Dispositional optimism, situational expectations, and the placebo response. *Journal of Psychosomatic Research*, 58(2), 121-127.
- Geers, A. L., Helfer, S. G., Weiland, P. E., & Kosbab, K. (2006). Expectations and placebo response: A laboratory investigation into the role of somatic focus. *Journal of Behavioral Medicine*, 29(2), 171-178.
- Geers, A. L., Kosbab, K., Helfer, S. G., Weiland, P. E., & Wellman, J. A. (2007). Further evidence for individual differences in placebo responding: An interactionist perspective. *Journal of Psychosomatic Research*, 62(5), 563-570.
- Geers, A. L., Weiland, P. E., Kosbab, K., Landry, S. J., & Helfer, S. G. (2005). Goal activation, expectations, and the placebo effect. *Journal of Personality and Social Psychology*, 89(2), 143-159.
- Gelman, A., & Hill, J. (2007). Data analysis using regression and

- multilevel/hierarchical models.
- Gelman, A., & Shalizi, C. (2010). Philosophy and the practice of bayesian statistics in the social sciences. *Oxford Handbook of the Philosophy of the Social Sciences*.
- Giebel, G. D., & Groeben, N. (2008). Social desirability in the measuring of patient satisfaction after treatment of coloproctologic disorders: on shortcomings of general bipolar satisfaction scales for quality management. *Langenbecks Arch Surg*, 393(4), 513-20.
- Gigerenzer, G., & Gaissmaier, W. (2011, January). Heuristic decision making. *Annual review of psychology*, 62, 451–82. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/21126183> doi: 10.1146/annurev-psych-120709-145346
- Giluk, T. (2009). Mindfulness, big five personality, and affect: A meta-analysis. *Personality and Individual Differences*, 47(8), 805–811.
- Goffaux, P., Redmond, W. J., Rainville, P., & Marchand, S. (2007). Descending analgesia - when the spine echoes what the brain expects. *Pain*, 130(1-2), 137-143.
- Gotzsche, P. C. (1995). Placebo effects. concept of placebo should be discarded. *BMJ*, 311(7020), 1640-1.
- Gracely, R., Dubner, R., Deeter, W., & Wolskee, P. (1985). Clinicians' expectations influence placebo analgesia. *The Lancet*, 325(8419), 43-43.
- Greenwald, & Banaji, M. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review*, 102(1), 4-26.
- Greenwald, Banaji, M. R., Rudman, L. A., Farnham, S. D., Nosek, B. A., & Mellott, D. S. (2002). A unified theory of implicit attitudes, stereotypes, self-esteem, and self-concept. *Psychological Review*, 109(1), 3-25.
- Greenwald, & Farnham, S. D. (2000). Using the implicit association test to measure self-esteem and self-concept. *Journal of Personality and Social Psychology*, 79(6), 1022-1038.
- Greenwald, McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, 74(6), 1464-1480.
- Greenwald, Nosek, B. A., & Banaji, M. R. (2003). 'understanding and using the implicit association test: I. an improved scoring algorithm': Correction to greenwald et al. (2003). *Journal of Personality and Social Psychology*, 85(3), 481-481.
- Greenwald, Poehlman, T. A., Uhlmann, E. L., & Banaji, M. R. (2009). Understanding and using the implicit association test : Iii. meta-analysis of predictive validity. *Journal of Personality and Social Psychology*, 97(1), 17-41.
- Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the implicit association test: I. an improved scoring algorithm. *Journal of Personality and Social Psychology*, 85(2), 197-216.

- Grossman, P., Niemann, L., Schmidt, S., & Walach, H. (2004). Mindfulness-based stress reduction and health benefits a meta-analysis. *Journal of Psychosomatic Research*, 57(1), 35-43.
- Grumm, M., Erbe, K., von Collani, G., & Nestler, S. (2008). Automatic processing of pain: The change of implicit pain associations after psychotherapy. *Behaviour Research and Therapy*, 46(6), 701-714.
- Grumm, M., & von Collani, G. (2007). Measuring big-five personality dimensions with the implicit association test—implicit personality traits or self-esteem? *Personality and Individual Differences*, 43(8), 2205-2217.
- Grünbaum, A. (1981). The placebo concept. *Behaviour research and therapy*, 19(2), 157-167.
- Gschwendner, T., Hofmann, W., & Schmitt, M. (2006). Synergistic moderator effects of situation and person factors of awareness and adjustment on the consistency of implicit and explicit attitudes. *Journal of Individual Differences*, 27(1), 47-56.
- Gschwendner, T., Hofmann, W., & Schmitt, M. (2008). Differential stability: The effects of acute and chronic construct accessibility on the temporal stability of the implicit association test. *Journal of Individual Differences*, 29(2), 70-79.
- Haberman, S. (1979). *Analysis of qualitative data*. Academic Press New York:.
- Hann, M., & Reeves, D. (2008). The sf-36 scales are not accurately summarised by independent physical and mental component scores. *Quality of Life Research*, 17(3), 11p.
- Hansen, E., Lundh, L.-G., Homman, A., & Wångby-Lundh, M. (2009). Measuring mindfulness: pilot studies with the swedish versions of the mindful attention awareness scale and the kentucky inventory of mindfulness skills. *Cognitive behaviour therapy*, 38(1), 2-15.
- Hays, R., Sherbourne, C., & Mazel, R. (1993). The rand 36-item health survey 1.0. *Health economics*, 2(3), 217-227.
- Heider, J. D., & Skowronski, J. J. (2007). Improving the predictive validity of the implicit association test. *North American Journal of Psychology*, 9(1), 53-76.
- Henrich, J., Heine, S. J., & Norenzayan, A. (2010). Most people are not weird. *Nature*, 466(7302), 29-29.
- Henson, R., & Roberts, J. (2006). Use of exploratory factor analysis in published research. *Educational and Psychological Measurement*, 66(3), 393.
- Herrnstein, R. J. (1962). Placebo effect in rat. *Science*, 138(3541), 677-&.
- Herzberg, P. Y., Glaesmer, H., & Hoyer, J. (2006). Separating optimism and pessimism: A robust psychometric analysis of the revised life orientation test (lot-r). *Psychological Assessment*, 18(4), 6p.
- Hofmann, W., Gawronski, B., Gschwendner, T., Le, H., & Schmitt, M. (2005). A meta-analysis on the correlation between the implicit association test and explicit self-report measures. *Personality and Social Psychology Bulletin*,

- 31(10), 1369-1385.
- Hofmann, W., Gschwendner, T., Castelli, L., & Schmitt, M. (2008). Implicit and explicit attitudes and interracial interaction: The moderating role of situationally available control resources. *Group Processes & Intergroup Relations*, 11(1), 69-87.
- Hofmann, W., & Schmitt, M. (2008). Advances and challenges in the indirect measurement of individual differences at age 10 of the implicit association test. *European Journal of Psychological Assessment*, 24(4), 207-209.
- Horn, J. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, 30(2), 179-185.
- Horne, R., Weinman, J., & Hankins, M. (1999). The beliefs about medicines questionnaire: The development and evaluation of a new method for assessing the cognitive representation of medication. *Psychology & Health*, 14(1), 1-24.
- Hrobjartsson, A. (1996). The uncontrollable placebo effect. *European journal of clinical pharmacology*, 50(5), 345-348.
- Hrobjartsson, A., & Gotzsche, P. C. (2001). Is the placebo powerless? an analysis of clinical trials comparing placebo with no treatment. *N Engl J Med*, 344(21), 1594-602.
- Hrobjartsson, A., & Gotzsche, P. C. (2003). Unreliable analysis of placebo analgesia in trials of placebo pain mechanisms. *Pain*, 104(3), 714 (1 pages).
- Hyland, M. E., Geraghty, A. W. A., Joy, O. E. T., & Turner, S. I. (2006). Spirituality predicts outcome independently of expectancy following flower essence self-treatment. *Journal of Psychosomatic Research*, 60(1), 53-58.
- Hyland, M. E., Whalley, B., & Geraghty, A. W. A. (2007). Dispositional predictors of placebo responding: A motivational interpretation of flower essence and gratitude therapy. *Journal of Psychosomatic Research*, 62(3), 331-340.
- Jensen, M. P., & Karoly, P. (1991). Motivation and expectancy factors in symptom perception: A laboratory study of the placebo effect. *Psychosomatic Medicine*, 53(2), 144-152.
- Joreskog, K. (1978). Structural analysis of covariance and correlation matrices. *Psychometrika*, 43(4), 443-477.
- Kabat-Zinn, J. (2003). Mindfulness-based interventions in context: Past, present, and future. *Clinical Psychology: Science & Practice*, 10(2), 13p.
- Kahler, C. W., Daughters, S. B., Leventhal, A. M., Gwaltney, C. J., & Palfai, T. P. (2007). Implicit associations between smoking and social consequences among smokers in cessation treatment. *Behaviour Research and Therapy*, 45(9), 2066-2077.
- Kahneman, D., Frederick, S., Gilovich, T., Griffin, D., Hertwig, R., Hilton, D., ... Frederick, S. (2002). Heuristics of Intuitive Judgment: Extensions and Applications. *System*, 1-30.
- Kaptchuk. (1998). Powerful placebo: The dark side of the randomised controlled

- trial. *Lancet*, 351(9117), 1722.
- Kaptchuk, Kelley, J. M., Conboy, L. A., Davis, R. B., Kerr, C. E., Jacobson, E. E., . . . Lembo, A. J. (2008). Components of placebo effect: randomised controlled trial in patients with irritable bowel syndrome. *BMJ: British Medical Journal*, 336(7651), 999-1003.
- Kaptchuk, Kelley, J. M., Deykin, A., Wayne, P. M., Lasagna, L. C., Epstein, I. O., . . . Wechsler, M. E. (2008). Do "placebo responders" exist? *Contemporary Clinical Trials*, 29(4), 587-595.
- Kaptchuk, T., Friedlander, Kelley, J., Sanchez, M., Kokkotou, E., Singer, J., . . . Lembo, A. (2010). Placebos without deception: a randomized controlled trial in irritable bowel syndrome. *PLoS One*, 5(12), e15591.
- Karpinski, A., & Steinman, R. B. (2006). The single category implicit association test as a measure of implicit social cognition. *Journal of Personality and Social Psychology*, 91(1), 16-32.
- Karpinski, A., Steinman, R. B., & Hilton, J. L. (2005). Attitude importance as a moderator of the relationship between implicit and explicit attitude measures. *Personality and Social Psychology Bulletin*, 31(7), 949-962.
- Kaufman, D. W., Kelly, J. P., Rosenberg, L., Anderson, T. E., & Mitchell, A. A. (2002). Recent patterns of medication use in the ambulatory adult population of the united states. *JAMA: the journal of the American Medical Association*, 287(3), 337-344.
- Kelly, G. (2003). *The psychology of personal constructs: Volume two: Clinical diagnosis and psychotherapy* (Vol. 2). Routledge.
- Kelly, G. A. (1991). *The psychology of personal constructs*. London ; New York :: Routledge in association with the Centre for Personal Construct Psychology.
- Kienle, G. S., & Kiene, H. (1997). The powerful placebo effect: Fact or fiction? *Journal of Clinical Epidemiology*, 50(12), 1311-1318.
- Kienle, G. S., & Kiene, H. (1998). The placebo effect: a scientific critique. *Complementary Therapies in Medicine*, 6(1), 14-24.
- Kirsch. (1985). Response expectancy as a determinant of experience and behavior. *American Psychologist*, 40(11), 1189-1202.
- Kirsch. (1994). Clinical hypnosis as a nondeceptive placebo: empirically derived techniques. *Am J Clin Hypn*, 37(2), 95-106.
- Kirsch. (1997). Response expectancy theory and application: A decennial review. *Applied and Preventive Psychology*, 6(2), 69-79.
- Kirsch. (1999). Hypnosis and placebos: Response expectancy as a mediator of suggestion effects. *Anales de Psicologia*, 15(1), 99-110.
- Kirsch, Moore, T. J., Scoboria, A., & Nicholls, S. S. (2002). The emperor's new drugs: An analysis of antidepressant medication data submitted to the u.s. food and drug administration. *Prevention & Treatment*, 5(1), 23.
- Kirsch, & Sapirstein, G. (1998). Listening to prozac but hearing placebo: A

- meta-analysis of antidepressant medication. *Prevention & Treatment*, 1(2), 2a.
- Kirsch, & Weixel, L. (1988). Double-blind versus deceptive administration of a placebo. *Behavioral Neuroscience*, 102(2), 319.
- Klauer, K. C., & Mierke, J. (2005). Task-set inertia, attitude accessibility, and compatibility-order effects: New evidence for a task-set switching account of the implicit association test effect. *Personality and Social Psychology Bulletin*, 31(2), 208-217.
- Klauer, K. C., Voss, A., Schmitz, F., & Teige-Mocigemba, S. (2007). Process components of the implicit association test: A diffusion-model analysis. *Journal of Personality and Social Psychology*, 93(3), 353-368.
- Kleijnen, J., De Craen, A. J. M., Van Everdingen, J., & Krol, L. (1994). Placebo effect in double-blind clinical trials: a review of interactions with medications. *Lancet(British edition)*, 344(8933), 1347-1349.
- Klinger, R., Soost, S., Flor, H., & Worm, M. (2007). Classical conditioning and expectancy in placebo hypoalgesia: A randomized controlled study in patients with atopic dermatitis and persons with healthy skin. *Pain*, 128(1-2), 31-39.
- Knipschild, P., & Arntz, A. (2005). Pain patients in a randomized trial did not show a significant effect of a positive consultation. *Journal of Clinical Epidemiology*, 58(7), 708-713.
- Kohls, N., Sauer, S., & Walach, H. (2009). Facets of mindfulness-results of an online study investigating the freiburg mindfulness inventory. *Personality and Individual Differences*, 46(2), 224-230.
- Lam, S.-f., Chiu, C.-y., & Lau, I. Y.-m. (2007). What do we learn from the implicit association test about intergroup attitudes in hong kong? the case of social identification inclusiveness and need for closure. *Asian Journal of Social Psychology*, 10(3), 123-130.
- Ledesma, D., & Kumano, H. (2009). Mindfulness-based stress reduction and cancer: a meta-analysis. *Psycho-Oncology*, 18(6), 9p.
- Levesque, C., & Brown, K. W. (2007). Mindfulness as a moderator of the effect of implicit motivational self-concept on day-to-day behavioral motivation. *Motivation and Emotion*, 31(4), 284-299.
- Levine, J. D., Gordon, N. C., & Fields, H. L. (1978). The mechanism of placebo analgesia. *Lancet*, 2(8091), 654-7.
- Linde, K., Witt, C. M., Streng, A., Weidenhammer, W., Wagenpfeil, S., Brinkhaus, B., ... Melchart, D. (2007). The impact of patient expectations on outcomes in four randomized controlled trials of acupuncture in patients with chronic pain. *Pain*, 128(3), 264-71.
- Little, R., & Rubin, D. (1987). Statistical analysis with missing data.
- Lord, F., Novick, M., & Birnbaum, A. (1968). *Statistical theories of mental test scores*. Addison-Wesley.
- MacCallum, R., & Austin, J. (2000). Applications of structural equation modeling in

- psychological research. *Annual review of psychology*, 51(1), 201–226.
- Macedo, A., Farrao, M., & Baatos, J.-E. (2003). Placebo effect and placebos: what are we talking about? some conceptual and historical considerations. *European Journal of Clinical Pharmacology*, 59(4), 337–342.
- MacKillop, J., & Anderson, E. J. (2007). Further psychometric validation of the mindful attention awareness scale (maas). *Journal of Psychopathology and Behavioral Assessment*, 29(4), 289–293.
- Mair, P., Hatzinger, R., & Maier, M. (2010). erm: Extended rasch modeling. [Computer software manual]. Retrieved from <http://CRAN.R-project.org/package=eRm> (R package version 0.12-0)
- Mårdby, A.-C., Åkerlind, I., & Jørgensen, T. (2007). Beliefs about medicines and self-reported adherence among pharmacy clients. *Patient education and counseling*, 69(1), 158–164.
- Martin-Pichora, A. L., Mankovsky-Arnold, T. D., & Katz, J. (2011). Implicit versus explicit associative learning and experimentally induced placebo hypoalgesia. *Journal of pain research*, 4, 67.
- McConnell, A. R., & Leibold, J. M. (2001). Relations among the implicit association test, discriminatory behavior, and explicit measures of racial attitudes. *Journal of Experimental Social Psychology*, 37(5), 435–442.
- McCullagh, P., & Nelder, J. (1989). *Generalized linear models (monographs on statistics and applied probability 37)*. Chapman Hall, London.
- McEachrane-Gross, F. P., Liebschutz, J. M., & Berlowitz, D. (2006). Use of selected complementary and alternative medicine (cam) treatments in veterans with cancer or chronic pain: a cross-sectional survey. *BMC complementary and alternative medicine*, 6(1), 34.
- McQuay, H., Carroll, D., & Moore, A. (1996). Variation in the placebo effect in randomised controlled trials of analgesics: all is as blind as it seems. *Pain*, 64(2), 331–335.
- Meagher, B. E., & Aidman, E. V. (2004). Individual differences in implicit and declared self-esteem as predictors of response to negative performance evaluation: Validating implicit association test as a measure of self-attitudes. *International Journal of Testing*, 4(1), 19–42.
- Meissner, K., Distel, H., & Mitzdorf, U. (2007). Evidence for placebo effects on physical but not on biochemical outcome parameters: a review of clinical trials. *BMC Med*, 5, 3.
- Micceri, T. (1989). The unicorn, the normal curve, and other improbable creatures. *Psychological Bulletin*, 105(1), 156.
- Mierke, J., & Klauer, K. C. (2003). Method-specific variance in the implicit association test. *Journal of Personality and Social Psychology*, 85(6), 1180–1192.
- Miller, F. G. (2008). Deceptive research. *Arch Gen Psychiatry*, 65(10), 1225–6;

- author reply 1226.
- Miller, F. G., & Kaptchuk, T. J. (2008). Deception of subjects in neuroscience: An ethical analysis. *Journal of Neuroscience*, 28(19), 4841-4843.
- Montgomery, David, D., DiLorenzo, T., & Schnur, J. (2007). Response expectancies and irrational beliefs predict exam-related distress. *Journal of Rational-Emotive & Cognitive-Behavior Therapy*, 25(1), 17-34.
- Montgomery, & Kirsch, I. (1997). Classical conditioning and the placebo effect. *Pain*, 72(1-2), 107-113.
- Moore, P., Duncan, G., Scott, D., Gregg, J., & Ghia, J. (1979). The submaximal effort tourniquet test: its use in evaluating experimental and chronic pain. *Pain*, 6(3), 375-382.
- Morris, L. A., & O'Neal, E. C. (1974). Drug-name familiarity and the placebo effect. *J Clin Psychol*, 30(3), 280-2.
- Morton, & Torgerson, D. J. (2003). Effect of regression to the mean on decision making in health care. *BMJ: British Medical Journal*, 326(7398), 1083-1084.
- Morton, Watson, A., El-Deredy, W., & Jones, A. (2009). Reproducibility of placebo analgesia: Effect of dispositional optimism. *Pain*, 146(1-2), 194.
- Myers, S., Phillips, R., Davis, R., Cherkin, D., Legedza, A., Kaptchuk, T., ... others (2008). Patient expectations as predictors of outcome in patients with acute low back pain. *Journal of general internal medicine*, 23(2), 148-153.
- Nes, L. S., & Segerstrom, S. C. (2006). Dispositional optimism and coping: A meta-analytic review. *Personality and social psychology review*, 10(3), 235-251.
- Nisbett, R., & Wilson, T. (1977). The halo effect: Evidence for unconscious alteration of judgments. *Journal of Personality and Social Psychology*, 35(4), 250-256.
- Nock, M. K., Ferriter, C., & Holmberg, E. (2007). Parent beliefs about treatment credibility and effectiveness: Assessment and relation to subsequent treatment participation. *Journal of Child and Family Studies*, 16(1), 27-38.
- Nosek, B. A., Greenwald, A. G., & Banaji, M. R. (2005). Understanding and using the implicit association test: II. method variables and construct validity. *Personality and Social Psychology Bulletin*, 31(2), 166-180.
- Nosek, B. A., & Smyth, F. L. (2007). A multitrait-multimethod validation of the implicit association test: Implicit and explicit attitudes are related but distinct constructs. *Experimental Psychology*, 54(1), 14-29.
- Ohayon, M. M., & Schatzberg, A. F. (2010). Chronic pain and major depressive disorder in the general population. *Journal of Psychiatric Research*, 44(7), 454-461.
- Oken, B. S., Flegal, K., Zajdel, D., Kishiyama, S., Haas, M., & Peters, D. (2008). Expectancy effect: Impact of pill administration on cognitive performance in healthy seniors. *Journal of Clinical & Experimental Neuropsychology*, 30(1), 7-17.
- Pearce, J. (1987). A model for stimulus generalization in pavlovian conditioning.

- Psychological Review*, 94(1), 61.
- Pearl, J. (1998). Graphs, causality, and structural equation models. *Sociological Methods & Research*, 27(2), 226–284.
- Perugini, M. (2005). Predictive models of implicit and explicit attitudes. *British Journal of Social Psychology*, 44(1), 29–45.
- Ploghaus, A., Becerra, L., Borras, C., & Borsook, D. (2003). Neural circuitry underlying pain modulation: expectation, hypnosis, placebo. *Trends in Cognitive Sciences*, 7(5), 197–200.
- Pollo, A., & Benedetti, F. (2009). Pain and the placebo effect. In *Biobehavioral approaches to pain* (p. 65–84).
- Popper, K. R. (1954). *Conjectures and refutations*. Minumsa.
- Praissman, S. (2008). Mindfulness-based stress reduction: A literature review and clinician's guide. *Journal of the American Academy of Nurse Practitioners*, 20(4), 5p.
- Price, D. D., Finniss, D. G., & Benedetti, F. (2008). A comprehensive review of the placebo effect: Recent advances and current thought. *Annual Review of Psychology*, 59(1), 565–590.
- Pruett, S. R., & Chan, F. (2006). The development and psychometric validation of the disability attitude implicit association test. *Rehabilitation Psychology*, 51(3), 202–213.
- Rasmussen, H., Scheier, M., & Greenhouse, J. (2009). Optimism and physical health: A meta-analytic review. *Annals of Behavioral Medicine*, 37(3), 239–256.
- Raz, A. (2007). Suggestibility and hypnotizability: Mind the gap. *American Journal of Clinical Hypnosis*, 49(3), 205–210.
- Revelle, W. (2015). psych: Procedures for psychological, psychometric, and personality research [Computer software manual]. Evanston, Illinois. Retrieved from <http://CRAN.R-project.org/package=psych> (R package version 1.5.1)
- Revelle, W., & Rocklin, T. (1979). Very simple structure: An alternative procedure for estimating the optimal number of interpretable factors. *Multivariate Behavioral Research*, 14(4), 403–414.
- Richetin, J., Perugini, M., Prestwich, A., & O'Gorman, R. (2007). The iat as a predictor of food choice: The case of fruits versus snacks. *International Journal of Psychology*, 42(3), 166–173.
- Rigdon, E. (1996). Cfi versus rmsea: A comparison of two fit indexes for structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, 3(4), 369–379.
- Rizopoulos, D. (2006). ltm: An r package for latent variable modelling and item response theory analyses. *Journal of Statistical Software*, 17(5), 1–25. Retrieved from <http://www.jstatsoft.org/v17/i05/>
- Robinson, F. P., Mathews, H. L., & Witek-Janusek, L. (2003). Psycho-endocrine-immune response to mindfulness-based stress reduction in

- individuals infected with the human immunodeficiency virus: A quasiexperimental study. *Journal of Alternative & Complementary Medicine*, 9(5), 12p.
- Roediger, H. L. (2013). Psychology's woes and a partial cure: The value of replication. *Psychology*.
- Rosenthal, R. (1967). Covert communication in the psychological experiment. *Psychological Bulletin*, 67(5), 356.
- Rosenthal, R. (1969). Interpersonal expectations: Effects of the experimenter's hypothesis. *Artifact in behavioral research*, 181, 277.
- Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, 48(2), 1–36. Retrieved from <http://www.jstatsoft.org/v48/i02/>
- Rudman, L., & Glick, P. (1999). Feminized management and backlash toward agentic women: The hidden costs to women of a kinder, gentler image of middle managers. *Journal of Personality and Social Psychology*, 77(5), 1004-1010.
- Rudolph, A., Schroder-Abe, M., Schutz, A., Gregg, A. P., & Sedikides, C. (2008). Through a glass, less darkly? reassessing convergent and discriminant validity in measures of implicit self-esteem. *European Journal of Psychological Assessment*, 24(4), 273-281.
- Ruth, A. B. (2006). Using self-report assessment methods to explore facets of mindfulness. *Assessment*, 13(1), 19p.
- Ruth, A. B., Gregory, T. S., & Kristin, B. A. (2004). Assessment of mindfulness by self-report: The kentucky inventory of mindfulness skills. *Assessment*, 11(3), 16p.
- Sass, D., & Schmitt, T. (2010). A comparative investigation of rotation criteria within exploratory factor analysis. *Multivariate Behavioral Research*, 45(1), 73–103.
- Sauro, M. D., & Greenberg, R. P. (2005). Endogenous opiates and the placebo effect: A meta-analytic review. *Journal of Psychosomatic Research*, 58(2), 115-120.
- Scheier, M. F., Carver, C. S., & Bridges, M. W. (1994). Distinguishing optimism from neuroticism (and trait anxiety, self-mastery, and self-esteem): a reevaluation of the life orientation test. *J Pers Soc Psychol*, 67(6), 1063-78.
- Schmukle, S. C., Back, M. D., & Egloff, B. (2008). Validity of the five-factor model for the implicit self-concept of personality. *European Journal of Psychological Assessment*, 24(4), 263-272.
- Schwarz, G. (1978). Estimating the dimension of a model. *The annals of statistics*, 6(2), 461–464.
- Schwarz, N. (1999). Self-reports: how the questions shape the answers. *American psychologist*, 54(2), 93.
- Scott, D. J., Stohler, C. S., Egnatuk, C. M., Wang, H., Koeppe, R. A., & Zubieta, J.-K. (2007). Individual differences in reward responding explain placebo-induced expectations and effects. *Neuron*, 55(2), 325-336.

- Segerstrom, S. C., Castaneda, J. O., & Spencer, T. E. (2003). Optimism effects on cellular immunity: testing the affective and persistence models. *Personality and Individual Differences*, 35(7), 1615-1624.
- Shapiro, A. K., & Shapiro, E. (1997). *The powerful placebo : from ancient priest to modern physician*. Baltimore :: Johns Hopkins University Press.
- Sherman, R., & Hickner, J. (2008). Placebos: current clinical realities. *J Clin Ethics*, 19(1), 62-5.
- Shiv, B., Carmon, Z., & Ariely, D. (2005). Placebo effects of marketing actions: Consumers may get what they pay for. *Journal of Marketing Research*, 42(4), 383-393.
- Singer, C., Grossman, I., Avidan, N., Beckmann, J. S., & Pe'er, I. (2005). Trick or treat: the effect of placebo on the power of pharmacogenetic association studies. *Hum Genomics*, 2(1), 28-38.
- Steffens, M. C., Kirschbaum, M., & Glados, P. (2008). Avoiding stimulus confounds in implicit association tests by using the concepts as stimuli. *British Journal of Social Psychology*, 47(2), 217-243.
- Steffens, M. C., & Konig, S. S. (2006). Predicting spontaneous big five behavior with implicit association tests. *European Journal of Psychological Assessment*, 22(1), 13-20.
- Sterzer, P., Frith, C., & Petrovic, P. (2008). Believing is seeing: expectations alter visual awareness. *Current Biology*, 18(16), R697-R698.
- Stewart-Williams, S. (2004). The placebo puzzle: Putting together the pieces. *Health Psychology*, 23(2), 198-206.
- Stewart-Williams, S., & Podd, J. (2004). The placebo effect: Dissolving the expectancy versus conditioning debate. *Psychological Bulletin*, 130(2), 324-340.
- Stigler, S. (1986). The history of statistics: The measurement of uncertainty before 1900.
- Stone, D. A., Kerr, C. E., Jacobson, E., Conboy, L. A., & Kaptchuk, T. J. (2005). Patient expectations in placebo-controlled randomized clinical trials. *J Eval Clin Pract*, 11(1), 77-84.
- Thomas, K. B. (1994). The placebo in general practice. *Lancet*, 344(8929), 1066-7.
- Thompson, Ritenbaugh, C., & Nichter, M. (2009). Reconsidering the placebo response from a broad anthropological perspective. *Culture, Medicine and Psychiatry*.
- Thompson, & Waltz, J. (2007). Everyday mindfulness and mindfulness meditation: Overlapping constructs or not? *Personality and Individual Differences*, 43(7), 1875-1885.
- Tindle, H. A., Chang, Y. F., Kuller, L. H., Manson, J. E., Robinson, J. G., Rosal, M. C., ... Matthews, K. A. (2009). Optimism, cynical hostility, and incident coronary heart disease and mortality in the women's health initiative. *Circulation*, 120(8), 656-62.
- Turner, J. A., & Deyo, R. A. (1994). *The importance of placebo effects in pain*

- treatment and research* (Vol. 271).
- Van Der Ark, L. (2007). Mokken scale analysis in r. *Journal of Statistical Software*, 20(11), 1–19.
- Van der Linden, W., & Hambleton, R. (1997). *Handbook of modern item response theory*. Springer New York.
- van Deventer, M. O. (2008). Meta-placebo: Do doctors have to lie about giving a fake treatment? *Medical Hypotheses*, 71(3), 335–339.
- Vase, L., Riley, J. L., & Price, D. D. (2002). A comparison of placebo effects in clinical analgesic trials versus studies of placebo analgesia. *Pain*, 99(3), 443–452.
- Vase, L., Robinson, M. E., Verne, G. N., & Price, D. D. (2005). Increased placebo analgesia over time in irritable bowel syndrome (ibs) patients is associated with desire and expectation but not endogenous opioid mechanisms. *Pain*, 115(3), 338–347.
- Venables, W., & Ripley, B. (2002). *Modern applied statistics with s*. Springer verlag.
- Vickers, A. J., & de Craen, A. J. M. (2000). Why use placebos in clinical trials? a narrative review of the methodological literature. *Journal of Clinical Epidemiology*, 53(2), 157–161.
- Vodouris, N. J., Peck, C. L., & Coleman, G. (1989). Conditioned response models of placebo phenomena further support. *Pain*, 38(1), 109–116.
- Voudouris, N. J., Peck, C. L., & Coleman, G. (1985). Conditioned placebo responses. *Journal of Personality and Social Psychology*, 48(1), 47–53.
- Voudouris, N. J., Peck, C. L., & Coleman, G. (1989). Conditioned response models of placebo phenomena: further support. *Pain*, 38(1), 109–116.
- Waber, R. L., Shiv, B., Carmon, Z., & Ariely, D. (2008). Commercial features of placebo and therapeutic efficacy. *JAMA*, 299(9), 1016.
- Walach, H., Schmidt, S., Dirhold, T., & Nosch, S. (2002). The effects of a caffeine placebo and suggestion on blood pressure, heart rate, well-being and cognitive performance. *International Journal of Psychophysiology*, 43(3), 247–260.
- Walker, S. S., & Schimmack, U. (2008). Validity of a happiness implicit association test as a measure of subjective well-being. *Journal of Research in Personality*, 42(2), 490–497.
- Wampol, B. E., Imel, Z. E., & Minami, T. (2007). The story of placebo effects in medicine: Evidence in context. *Journal of Clinical Psychology*, 63(4), 379–390.
- Watson, A., El-Deredy, W., Bentley, D. E., Vogt, B. A., & Jones, A. K. (2006). Categories of placebo response in the absence of site-specific expectation of analgesia. *Pain*, 126(1–3), 115–22.
- Weber, S., & Cook, T. (1972). Subject effects in laboratory research: An examination of subject roles, demand characteristics, and valid inference. *Psychological Bulletin*, 77(4), 273.
- Whalley, B., Hyland, M. E., & Kirsch, I. (2008). Consistency of the placebo effect. *Journal of Psychosomatic Research*, 64(5), 537–541.

- Wickramasekera, I. (2001). The placebo efficacy study: Problems with the definition of the placebo and mechanisms of placebo efficacy. *Advances in Mind-Body Medicine*, 17(4), 309-312.
- Wittenbrink, B., & Schwarz, N. (2007a). *Implicit measures of attitudes*. New York, NY US: Guilford Press.
- Wittenbrink, B., & Schwarz, N. (2007b). Measuring attitudes through priming. In *Implicit measures of attitudes*. (p. 17-58). New York, NY US: Guilford Press.
- Zwick, W., & Velicer, W. (1986). Comparison of five rules for determining the number of components to retain. *Psychological Bulletin*, 99(3), 432-442.