DISABILITY AND PHYSICAL ACTIVITY BEHAVIOURS: AN APPLICATION OF THEORETICAL FRAMEWORKS

NICOLA HOBBS

FEBRUARY 2010

A thesis submitted for the degree of Doctor of Philosophy

University of Stirling

ACKNOWLEDGEMENTS

I would first like to thank Diane Dixon as my principal supervisor, who has given me every possible opportunity to explore and develop my ideas and research questions. I am extremely lucky to have been supervised by Diane and I thank her for her continual support and encouragement over the last three years.

I am also very grateful to Ronan O’Carroll as my second supervisor for his support and advice.

I have also received advice and assistance from many other people whilst conducting the studies reported in this thesis and would particularly like to thank Paul Dieppe and the ‘EUROHIP’ consortium, Marie Johnston, Kate Howie, Susan Rasmussen, Tony Little, Stuart Galloway, Angus Hunter, the University fitness centre officers and the Stirling Health and Social Research group. I also owe many thanks to study participants.

I would like to thank many friends for their support including Ben, Cath, Cass and Bristol pals. También quiero agradecer a mis amigos de Guatemala, específicamente Luz María y los demás de la ODHAG, y “las chulas”.

I would also like to thank friends I’ve made in Scotland, especially Roser, Ann, Maggie, Janet, office mates and gym mates. Plus, all my scuba diving buddies over the last few years who have kept my head safely under the water!

Finally, I owe many thanks to my family for their love and support. Stacey, who reminds me not to unnecessarily worry and that things work out for the best in the end; Joey, who has imparted her wisdom as someone who has been here before; and Dad and Anne who listen, advise and support whatever the outcome.

ABSTRACT

Background: The prevalence of disability increases with age; therefore with an aging population, interventions to reduce disability are crucial. This thesis adopts a behavioural conceptualisation of disability. The theoretical frameworks of the International Classification of Functioning, Disability and Health (ICF), the Theory of Planned Behaviour (TPB) and the integrated ICF/TPB model are applied to investigate disability and physical activity (PA) behaviours. The thesis aims to: (1) identify the factors involved in the prioritisation of patients for total joint replacement; (2) classify patient pre-operative expectations of total hip replacement (THR) and investigate the relationship between expectations and recovery after surgery, and; (3) test whether the TPB and theory-based interventions can predict and explain PA *within* individuals. **Method:** Five studies were conducted. In the first study, health professionals judged whether the items from two prioritisation tools measured each of the ICF constructs. In the second study, surgeons ranked patient vignettes, which differed by constructs from the integrated model, in order of priority for THR. In the third study, a large cohort of THR patients reported expectations of surgery pre-operatively. Health and functioning were also reported pre-operatively and 1-year post-operatively. The fourth and fifth studies were a series of experimental *n-of-1* studies using diary methods assessing TPB cognitions and PA behaviours. **Results:** There is a lack of agreement between judges in relation to the content of many of the items from prioritisation tools. Behavioural and psychological factors can influence prioritisation for THR. The majority of patient expectations of THR addressed activities and social participation; however, the evidence for a relationship between expectations and recovery was limited. The TPB can predict PA *within* some individuals but the evidence in support of interventions to increase PA was limited. **Discussion:** The findings provide important clinical and theoretical implications for understanding disability and physical activity behaviours.

EXECUTIVE SUMMARY

Chapter 1 – General introduction

Disability can be conceptualised as behaviour. This thesis employed theoretical frameworks to investigate disability and physical activity behaviours. A broad introduction to theoretical frameworks applied to disability, the health condition of osteoarthritis and concurrent disability, and theory-based interventions to reduce disability and increase physical activity is presented.

Chapter 2 – Prioritisation for total joint replacement – operationalisation of clinical priority assessment tools

Total joint replacement surgery (TJR) can reduce pain and disability. However, provision does not always meet demand; therefore, patients have to be assigned priority on waiting lists for surgery. Clinical prioritisation tools have been developed to standardise this prioritisation decision. The method of discriminant content validation was used to determine whether each item from two prioritisation tools for TJR measured each of the theoretical constructs of the International Classification of Functioning, Disability and Health (ICF). A large number of items could not be classified as measuring any of the ICF constructs, indicating a lack of agreement between the expert judges. Variability in the interpretation of items limits the ability of prioritisation tools to standardise this clinical decision.

Chapter 3 – Prioritisation for total hip replacement – an exploratory investigation of the factors that influence prioritisation

The use of tools to aid the prioritisation of patients for TJR, such as those examined in Chapter 2, is not standard practice. Some clinicians employ his or her own clinical decision making. Conjoint analyses were used to examine the influence of clinical and psychological patient attributes on orthopaedic surgeons' prioritisation for total hip replacement (THR). The data from this study indicated that surgeons placed similar importance on a behavioural construct, which combined the ICF constructs of activity limitation and participation restriction, as on impairment (pain). Furthermore, when patient motivation is considered, it had a greater influence on prioritisation than either impairment or activity limitation. The evidence suggests that surgeons apply a behavioural model of disability that recognises both medical and psychological constructs in this decision making process.

Chapter 4 – Predicting recovery from total hip replacement – the role of pre-operative patient expectations

Patient expectations have been suggested to explain individual variability in surgical recovery and health outcomes. This study employed the ICF, as a model of health outcomes, to investigate the role of patient expectations on recovery after THR. A large European cohort of THR patients was studied prospectively. Support for the ICF as a suitable framework to classify patient expectations was obtained; all expectations were classified to the ICF constructs. THR targets impairment, yet the majority of expectations were classified as activity limitations and participation restrictions. Limited evidence for the role of patient expectations on recovery was found. However, limitations associated with the wording of the expectation questions are recognised, which may have induced response bias towards behavioural expectations.

Chapter 5 – The *n-of-1* methodology and experimental design

The background and rationale for the following series of experimental *n-of-1* studies, which test the ability of the Theory of Planned behaviour (TPB) to explain physical activity (PA) behaviour within individuals, is presented. The study methodology is also detailed.

Chapter 6 – Testing the ability of the integrated ICF/TPB model to explain physical activity behaviour within individuals with osteoarthritis: experimental *n-of-1* studies

In individuals with osteoarthritis, control cognitions have been shown to predict and explain activity limitations. Pain has also been shown to be associated with PA levels. The integrated ICF/TPB model was applied to explain PA behaviour within individuals with osteoarthritis. TPB cognitions, pain and PA were measured using diary methods for 12 weeks. At six weeks each participant received a theory-based intervention. There was support for the TPB to predict PA in all four participants. A weak relationship between pain and PA was also identified. The data provided some support for action planning but no support for the candidate causal pathway between self-efficacy and PA.

Chapter 7 – Testing the ability of the TPB to explain physical activity behaviour within healthy individuals: experimental *n-of-1* studies

The evidence in support of the TPB to explain PA is dominated by group-based designs. This study tested the ability of the TPB to predict PA behaviour and behaviour change within six healthy individuals. The same experimental design that was used in Chapter 6 was applied. There was support for the TPB to predict PA in three of six participants. There was evidence in support of the candidate causal pathway between perceived controllability and PA, but not between self-efficacy and PA. Action planning did not receive any support within individuals. The implications of the findings from Chapters 6 and 7 are discussed in relation to theory testing within individuals.

Chapter 8 – General discussion

A general discussion of the studies compiling this thesis is presented, including general limitations, implications, future work and anecdotal observations.

PAPERS

The following is a list of papers arising from this thesis:

Hobbs, N. & Dixon, D. (in preparation). Testing the International Classification of Functioning. Disability and Health and the Theory of Planned Behaviour to explain physical activity within individuals with mobility problems. *Rehabilitation Psychology (target journal).*

Hobbs, N. & Dixon, D. (in preparation). What do prioritisation tools for joint replacement actually measure? *Journal of Health Services Research & Policy (target journal).*

Hobbs, N. & Dixon, D. (in preparation). A conjoint analysis study of surgeons' prioritisation of patients for total hip replacement. *Social Science and Medicine (target journal).*

Hobbs, N. & Dixon, D., Dieppe, P., and the EUROHIP group (in preparation). Do patients have impairment, activity limitation or participation restriction expectations of hip replacement before surgery? *Health and Quality of Life Outcomes (target journal).*

Hobbs, N. & Dixon, D., Dieppe, P., and the EUROHIP group (in preparation). The role of patient pre-operative expectations on recovery after total hip replacement surgery. *Arthritis and Rheumatism (target journal).*

Hobbs, N., Dixon, D., Johnston, M. & Howie, K. (in preparation). Do theory-based interventions work within individuals? *Health Psychology (target journal).*

TABLE OF CONTENTS

CHAPTER 1 - GENERAL INTRODUCTION **1**

1.1 INTRODUCTION 1

1.2 DISABILITY AS BEHAVIOUR 2

1.3 THEORETICAL FRAMEWORKS APPLIED TO DISABILITY 3

1.3.1 The World Health Organisation International Classification of Functioning, Disability and Health (ICF) 3

1.3.1.1 The ICF taxonomy 5

1.3.1.2 ICF core measurement sets 5

1.3.1.3 Linking items from health status measures to the ICF: the problem of discriminant validity 6

1.3.2 Psychological Models: The Theory of Planned Behaviour 8

1.3.2.1 The Theory of Planned Behaviour (TPB) model 8

1.3.3.2 Applications of the TPB to health-related behaviours 9

1.3.3 Integrating the ICF and the TPB 10

1.4 OSTEOARTHRITIS AND DISABILITY 12

1.4.1 Osteoarthritis Symptoms and Consequences 13

1.4.1.1 Management of osteoarthritis 14

1.5 THEORY-BASED INTERVENTIONS 15

1.5.1 Interventions on Impairment to Reduce Disability: Total Joint Replacement 15

1.5.1.1 Indications for TJR 16

1.5.1.2 Prioritisation for TJR 18

1.5.2 Interventions on Cognitions to Reduce Disability and Increase PA 19

1.5.2.1 Limitations of group design studies 19

1.6 AIMS OF THIS THESIS 21

CHAPTER 2 - PRIORITISATION FOR TOTAL JOINT REPLACEMENT – OPERATIONALISATION OF CLINICAL PRIORITY ASSESSMENT TOOLS **22**

2.1 ABSTRACT 22

2.2 INTRODUCTION 24

2.2.1 Research Questions 28

2.3 METHODOLOGY 28

2.3.1 Design 28

2.3.2 Participants 29

2.3.3 Materials 29

2.3.4 Procedure 29

2.3.5 Analyses 30

2.3.6 Ethics 31

2.4 RESULTS 32

2.4.1 New Zealand Priority Criteria Project (NZPC) Tool 32
2.4.1.1 Inter-rater reliability 32
2.4.1.2 Discriminant content validation 32
2.4.2 Western Canada Waiting List Project (WCWL) Tool 33
2.4.2.1 Inter-rater reliability 33
2.4.2.2 Discriminant content validation 34
2.5 DISCUSSION 35
2.5.1 Summary and Implications 40

CHAPTER 3 - PRIORITISATION FOR TOTAL HIP REPLACEMENT – AN EXPLORATORY INVESTIGATION OF THE FACTORS THAT INFLUENCE PRIORITISATION 42

3.1 ABSTRACT 42
3.2 INTRODUCTION 44
3.2.1 Research Questions 48
3.3 METHODOLOGY 48
3.3.1 Design 48
3.3.2 Participants 48
3.3.3 Measures and Materials 49
3.3.4 Procedure 50
3.3.5 Analyses 51
3.3.6 Ethics 52
3.4 RESULTS 52
3.4.1 Set 1:Impairment (I), Activity Limitation (A) & Participation Restriction (P) 52
3.4.2 Set 2 – Impairment (I), Activity Limitation (A) & Patient Motivation 53
3.5 DISCUSSION 55
3.5.1 Summary and Implications 62

CHAPTER 4 - PREDICTING RECOVERY AFTER TOTAL HIP REPLACEMENT – THE ROLE OF PRE-OPERATIVE PATIENT EXPECTATIONS 63

4.1 ABSTRACT 63
4.2 INTRODUCTION 65
4.2.1 'EUROHIP' Consortium 70
4.2.2 Research Questions 71
4.3 METHODOLOGY 71
4.3.1 Design 71
4.3.2 Participants 72
4.3.3 Measures 73
4.3.3.1 Disability 73
4.3.3.2 Health status and health-related quality of life 73
4.3.3.3 Recovery 74
4.3.3.4 Patient expectations 75

4.3.3.5 Met or unmet patient expectations 77
4.3.3.6 Demographic and clinical variables 78
4.3.4 Procedure 79
4.3.5 Analyses 79
4.3.6 Ethics 82
4.4 RESULTS 82
4.4.1 Attrition 82
4.4.2 Description of Sample Before Surgery 83
4.4.3 Research Question 1: Description of Pre-Operative Patient Expectations 85
4.4.4 Correlation Analyses 86
4.4.5 Research Question 2: Number of Expectations Analyses 88
4.4.5.1 Predicting recovery from disability (WOMAC) by the number of 'need' ....expectations 88
4.4.5.2 Predicting recovery from health-related quality of life (EQ-5D) by the .... number of 'need' expectations 90
4.4.5.3 Predicting recovery from disability (WOMAC) by the number of 'desire' ....expectations 91
4.4.5.4 Predicting recovery from heath-related quality of life (EQ-5D) by the .... number of 'desire' expectations 93
4.4.5.5 Post-hoc analysis of the relationship between recovery and responding to the .... expectation questions 93
4.4.6 Research Question 3: Content of Expectations Analyses 94
4.4.6.1 Relationship between the content of 'need' expectations and recovery 94
4.4.6.2 Relationship between the content of 'desire' expectations and recovery 94
4.4.7 Research Question 4: Association Between Recovery and Having Expectations Met 97
4.4.8 Post-Hoc Power Analysis 100

4.5 DISCUSSION 100
4.5.1 Summary and Implications 110
CHAPTER 5 - THE *N-OF-1* METHODOLOGY AND EXPERIMENTAL DESIGN **111**
5.1 BACKGROUND AND RATIONALE 111
5.1.1 Current Use and Evidence in Support of the Theory of Planned Behaviour 111
5.1.2 The n-of-1 Experimental Methodology 116
5.1.3 Ecological Momentary Assessment, Diary Methods and the n-of-1 Methodology 118
5.1.4 Physical Activity Behaviour 120
5.1.5 General Research Questions Addressed in Chapters 6 and 7 122
5.2 DESIGN 122
5.3 PARTICIPANTS 123
5.4 MEASURES 123
5.4.1 Behaviour 123

5.4.2 TPB Cognitions 124
5.4.3 Health Status Measures 124
5.5 INTERVENTION TYPES 125
5.5.1 TPB-based Intervention Targeting Either PC or SE 125
5.5.2 Action Planning Intervention 125
5.6 PROCEDURE 126
5.7 INTERVENTION PROTOCOL 126
5.8 ANALYSES 128
5.8.1 Criteria to Evaluate Theory 130

CHAPTER 6 - TESTING THE ABILITY OF THE INTEGRATED MODEL TO EXPLAIN PHYSICAL ACTIVITY BEHAVIOUR WITHIN INDIVIDUALS WITH OSTEOARTHRITIS: EXPERIMENTAL *N-OF-1* STUDIES **131**

6.1 ABSTRACT 131
6.2 INTRODUCTION 133
6.2.1 Research Questions 135
6.3 METHODOLOGY 136
6.3.1 Design 136
6.3.2 Participants 136
6.3.3 Measures and Materials 137
6.3.3.1 PA behaviours 137
6.3.3.2 TPB cognitions 138
6.3.3.3 Other diary measures 138
6.3.3.4 Health status measures 138
6.3.3.5 Materials 139
6.3.4 Procedure 139
6.3.5 Intervention protocol 140
6.3.6 Analyses 140
6.3.7 Ethics 140
6.4 RESULTS 141
6.4.1 Description of Participants 141
6.4.2 Diary Completion 143
6.4.3 Overall Variability in TPB Cognitions and PA Behaviours 144
6.4.4 Baseline Data Analyses 145
6.4.4.1 TPB predictors of PA behaviours 145
6.4.4.2 Intervention selection 146
6.4.5 Effect of SE Intervention for Participant 1 147
6.4.5.1 Effect on SE and the chosen PA behaviour 147
6.4.6 Effect of Action Planning Intervention for Participants 2 and 4 148
6.4.6.1 Effect of action planning on the chosen PA behaviour 148
6.4.6.2 Effect of action planning on the intention-behaviour relationship 149

6.4.7 Analyses Relating to Participant 3 Who Declined the Action Planning Intervention 150
6.4.8 Relationship Between Pain and PA During the Baseline Period 151
6.5 DISCUSSION 152
6.5.1 Is There Support for the TPB to Predict PA Within Individuals with Osteoarthritis? 152
6.5.2 Is there Support for a Causal Pathway Between SE and PA Behaviour Within Individuals with Osteoarthritis? 156
6.5.3 Is there Support for Action Planning Within Individuals with Osteoarthritis? 156
6.5.4 What is the Relationship Between Pain and PA Within Individuals with Osteoarthritis? 159
6.5.5 Summary 162
CHAPTER 7 - TESTING THE ABILITY OF THE THEORY OF PLANNED BEHAVIOUR TO EXPLAIN PHYSICAL ACTIVITY WITHIN HEALTHY INDIVIDUALS: EXPERIMENTAL *N-OF-1* STUDIES **163**
7.1 ABSTRACT 163
7.2 INTRODUCTION 165
7.2.1 Research Questions 165
7.3 METHODOLOGY 166
7.3.1 Design 166
7.3.2 Participants 166
7.3.3 Measures and Materials 167
7.3.3.1 PA behaviours 167
7.3.3.2 TPB cognitions 168
7.3.3.3 Health status, fitness and physiology measures 168
7.3.3.4 Materials 169
7.3.4 Procedure 169
7.3.5 Intervention Protocol 169
7.3.6 Analyses 170
7.3.7 Ethics 170
7.4 RESULTS 170
7.4.1 Description of Participants 170
7.4.2 Diary Completion 173
7.4.3 Overall Variability in TPB Cognitions and PA Behaviours 174
7.4.4 Baseline Data Analyses 176
7.4.4.1 TPB predictors of PA behaviours 176
7.4.4.2 Intervention selection 177
7.4.5 Effect of PC or SE Intervention for Participants 1, 4 and 5 178
7.4.5.1 Effect of the individually tailored PC intervention on PC and the chosen PA .for participants 1 and 4 178
7.4.5.2 Effect of the default SE intervention on SE and the chosen PA - participant 5 180
7.4.6 Effect of Action Planning Intervention for Participants 2, 3 and 6 181

7.4.6.1 Effect of the individually tailored action planning intervention on the chosen PA and the intention-behaviour relationship for participant 6 181
7.4.6.2 Effect of the default action planning intervention on the chosen PA and the intention-behaviour relationship for participants 2 and 3 182
7.5 DISCUSSION 184
7.5.1 Is There Support for the TPB to Predict PA Within Healthy Individuals? 184
7.5.2 Individually Tailored Interventions 185
7.5.2.1 Is there support for a causal pathway between PC and PA behaviour within healthy individuals (participants 1 and 4)? 185
7.5.2.2 Is there support for action planning within healthy individuals (participant 6)? 187
7.5.3 Default Interventions 190
7.5.3.1 Is there support for a causal pathway between SE and PA behaviour within healthy individuals (participant 5)? 190
7.5.3.2 Is there support for action planning within healthy individuals (participants 2 and 3)? 191
7.5.4 Summary 193
7.5.5 Summary and Implications of the Findings of the Studies Reported in Chapters 6 & 7 194

CHAPTER 8 - GENERAL DISCUSSION 197

8.1 WHAT WAS KNOWN BEFORE THIS THESIS 197
8.2 WHAT THIS THESIS ADDS 197
8.3 GENERAL LIMITATIONS 198
8.4 IMPLICATIONS AND FUTURE WORK 199
8.5 ANECDOTAL OBSERVATIONS IN THE *N-OF-1* STUDIES 204
8.5.1 Participant Burden and Feedback on Participation 204
8.5.2 Self-Monitoring Effects and Active Components of the Behaviour Change Interventions 205

REFERENCES 208

APPENDIX 1 - NEW ZEALAND PRIORITY CRITERIA (NZPC) FOR JOINT REPLACEMENT 229

APPENDIX 2 - WESTERN CANADA WAITING LIST CRITERIA (WCWL) FOR JOINT REPLACEMENT 230

APPENDIX 3 - CONJOINT ANALYSIS STUDY PATIENT VIGNETTES 231

APPENDIX 4 - EXAMPLE MORNING DIARY USED IN CHAPTER 6 232

APPENDIX 5 - TIME PLOTS AND FURTHER ANALYSES OF INTERVENTION EFFECTS FOR CHAPTER 6 233

APPENDIX 6 - EXAMPLE DIARY USED IN CHAPTER 7 235

APPENDIX 7 - TIME PLOTS AND FURTHER ANALYSES OF INTERVENTION EFFECTS FOR CHAPTER 7 236

LIST OF TABLES

TABLE 2.1 Definitions of the WHO-ICF constructs 29

TABLE 2.2 Classification of NZPC items to the ICF 33

TABLE 2.3 Classification of WCWL items to the ICF 35

TABLE 3.1 The type and level of attributes expressed in each patient vignetter for each set 50

TABLE 3.2 Relative average importance values (%) for the attributes expressed in each set 53

TABLE 4.1 Definitions of the WHO-ICF constructs 77

TABLE 4.2 Demographic and clinical characteristics of the sample pre-operatively (n=1108) .. 84

TABLE 4.3 Proportion of expectations in each of the ICF construct classification groups for the first three patient 'need' expectations and the first three 'desire' expectations 85

TABLE 4.4 Correlation between demographic and clinical characteristics, patient expectations, and recovery variables 87

TABLE 4.5 Summary of results of multiple linear regressions of Total WOMAC recovery and WOMAC subscales on demographic and clinical characteristics and the number of 'need' expectations 89

TABLE 4.6 Summary of results of multiple linear regression of EQ-5D recovery on demographic and clinical characteristics and the number of 'need' expectations 90

TABLE 4.7 Summary of results of multiple linear regressions of Total WOMAC recovery and WOMAC subscales on demographic and clinical characteristics and the number of 'desire' expectations 92

TABLE 4.8 Summary of results of multiple linear regression of EQ-5D recovery on demographic and clinical characteristics and the number of 'desire' expectations 93

TABLE 4.9 Descriptive statistics for recovery variables by the ICF content of the first three 'need' expectations given by patients 95

TABLE 4.10 Descriptive statistics for recovery variables by the ICF content of the first three 'desire' expectations given by patients.. 96

TABLE 4.11 Proportion of expectations in each of the ICF construct classification groups for the first three 1-year post-operative current patient function responses 98

TABLE 4.12 Binary logistic regression predicting likelihood of having pre-operative expectations 'met' by Total WOMAC recovery and EQ-5D recovery – Model 1 99

TABLE 4.13 Binary logistic regression predicting likelihood of having pre-operative expectations 'met' by demographic and clinical variables and EQ-5D recovery- Model 2 100

TABLE 6.1 Pre- and post-study description of participants 142

TABLE 6.2 Baseline cross-correlations and time lags between TPB cognitions and the physical activity selected for intervention, and type of intervention given 146

TABLE 6.3 Effect of self-efficacy (SE) intervention for walking for participant 1 148

TABLE 6.4 Effect of action planning intervention for walking for participant 2 149
TABLE 6.5 Effect of action planning intervention for driving for participant 4 149
TABLE 6.6 Pre- and post-intervention data for participant 3 who declined the action planning intervention on walking 150
TABLE 6.7 Descriptive statistics of pain scores, cross-correlations and time lags between pain and physical activity (PA) behaviours at baseline 151
TABLE 7.1 Pre- and post-study description of participants 172
TABLE 7.2 Baseline cross-correlations and time lags between TPB cognitions and the physical activity selected for intervention, and type of intervention given 177
TABLE 7.3 Effect of perceived controllability (PC) intervention for participants 1 and 4........ 180
TABLE 7.4 Effect of self-efficacy (SE) intervention for walking for participant 5 181
TABLE 7.5 Effect of action planning intervention for gym behaviour for participant 6........... 182
TABLE 7.6 Effect of action planning intervention for participants 2 and 3............................. 183

LIST OF FIGURES

FIGURE 1.1 The WHO International Classification of Impairments, Disabilities and Handicaps (ICIDH) 2
FIGURE 1.2 The WHO International Classification of Functioning, Disability and Health (ICF) 4
FIGURE 1.3 The Theory of Planned Behaviour adapted from Ajzen, 2006 9
FIGURE 1.4 The integrated ICF/TPB model of disability 11
FIGURE 3.1 Mean part-worth utility values for moderate and severe levels of impairment (I), activity limitation (A) and participation restriction (P) 53
FIGURE 3.2 Mean part-worth utility values for moderate and severe levels of impairment (I) and activity limitation (A), and no and high levels of motivation (M) 54
FIGURE 6.1 Examples of time plots of variability in PC and SE cognitions for logging and driving behaviours in participants 3 and 4 respectively 145
FIGURE 7.1 Examples of time plots of variability in TPB cognitions for PA behaviours illustrating various types of variability: frequent fluctuations (participant 1), ceiling effects (participant 3), floor effects (participant 4) and extreme response set (participant 5).............. 176
FIGURE 7.2 Effect of perceived controllability (PC) intervention on the chosen behaviour of ‘going for a walk’ (measured by pedometer step count) for participant 1.............................. 179

INDEX OF ABBREVIATIONS

A: Activity limitation
ASA: American Society of Anesthesiology comorbidity score
BMI: Body Mass Index
CI: Confidence Intervals
DCV: Discriminant Content Validation
EQ-5D: EuroQol
HADS: Hospital Anxiety and Depression Scale
ICC: Intraclass Correlation Coefficient
ICF: International Classification of Functioning, Disability and Health
ICIDH: International Classification of Impairments, Disabilities and Handicaps
I: Impairment
LTEQ: Leisure Time Exercise Questionnaire
NZPC: New Zealand Priority Criteria Project
P: Participation restriction
PA: Physical Activity
PC: Perceived Controllability
SE: Self-Efficacy
THR: Total Hip Replacement
TJR: Total Joint Replacement
TKR: Total Knee Replacement
TPB: Theory of Planned Behaviour
WCWL: Western Canada Waiting List Project
WHO: World Health Organization
WOMAC: Western Ontario and McMaster Universities Osteoarthritis Index

Chapter 1 - General introduction

1.1 Introduction

The prevalence of disability in adults increases with age (Martin, Meltzer, & Elliott, 1988; World Health Organization, 2004). Therefore, with an aging population, the total number of older people who are functionally limited is also increasing (Boult, Altmann, Gilbertson et al., 1996). The average global prevalence of moderate and severe disability is 46% in people over 60 years (World Health Organization, 2004). Moreover, figures from England report that the prevalence of locomotor disability in people over 65 years is approximately 32% of people who live in private homes and approximately 76% of men and 81% of women who live in residential or nursing homes (Hirani & Malbut, 2002). This high, and increasing, prevalence of disability in the older adult population means that interventions to reduce disability are crucial.

The approach to the development of interventions to reduce disability differs depending on how disability is conceptualised. The medical conceptualisation of disability emphasises the pathology of disablement and therefore, an intervention to reduce disability based on the medical approach targets the biological impairment. In contrast, the behavioural conceptualisation of disability recognises biomedical, psychological and social factors and therefore, an intervention to reduce disability based on the behavioural approach is complex. The development and evaluation of complex interventions delivered by multidisciplinary health professional teams require an integrative theoretical model of disability that conceptualises disability as behaviour.

1.2 Disability as Behaviour

Early work in the area of disability was dominated by medical models of disability, which conceptualised disability as a simple consequence of a bodily impairment. The World Health Organization (WHO) International Classification of Impairments, Disabilities and Handicaps (ICIDH: World Health Organization, 1980) provided the theoretical framework for this medical conceptualisation of disability (see Figure 1.1) and the following definition of disability:

> *Disability is any restriction or lack (resulting from impairment) of ability to perform an activity in the manner or within the range considered normal for a human being*

Figure 1.1 The WHO International Classification of Impairments, Disabilities and Handicaps (ICIDH)

However, the ICIDH was subject to much empirical and theoretical criticism. For example, it was criticised for its conceptualisation of disability as a simple consequence of impairment and its failure to acknowledge psychological and social factors that influence disability (Bickenbach, Chatterji, Badley et al., 1999; Imrie, 2004). In response to these criticisms, the WHO worked to develop a more comprehensive model of disability which incorporates biomedical, social and psychological factors, and in 2001, the ICIDH was replaced by the International Classification of Functioning, Disability and Health (ICF: World Health Organization, 2001). Unlike its predecessor, the ICF does not adopt a medical conceptualisation of disability as a simple consequence of impairment, but rather adopts a behavioural approach viewing disability as a component of health. The conceptualisation of disability as behaviour encourages psychological explanations of disability. Therefore, psychological theories of behaviour and behaviour change that identify the role of cognitions in behaviour can also be applied to the study of disability (Johnston, 1997).

1.3 Theoretical Frameworks Applied to Disability

1.3.1 The World Health Organisation International Classification of Functioning, Disability and Health (ICF)

The ICF is an internationally accepted conceptual framework that is intended to be a theoretical and practical classification system to understand functioning and disability associated with health and illness (see Figure 1.2). Unlike the earlier WHO-ICIDH model, disability is no longer viewed as a consequence of disease (impairment) but rather a component of health. The ICF states that a health condition can be manifested at three levels: the biological, the individual and the societal. These manifestations or health outcomes comprise the core constructs of the model. The ICF framework has two parallel arms: one for functioning and one for disability. The core constructs in each arm are conceptualised differently. Within the context of functioning, the core constructs of the ICF are ‘body function & structure’, ‘activity’ and ‘social participation’, shown in the non-italicised text in Figure 1.2. ‘Body function & structure’ are the physiological and psychological functions of body systems, and the anatomical parts of the body. ‘Activity’ is the execution of a task or action by an individual. ‘Social participation’ is the involvement in life situations. Within the context of disability, the core constructs of the ICF are *impairment*, *activity limitation* and *participation restriction*, shown in the italicised text in Figure 1.2, and reflect variation in one or a combination of the functioning constructs. *Impairment* is a significant deviation or loss in body function or structure. *Activity limitations* are difficulties an individual may have in executing activities. *Participation restrictions* are problems an individual may experience in involvement in life situations.

Health Condition
disorder/disease

Body function & structure
impairment

Activity
activity limitation

Participation
participation restriction

Environmental
factors

Personal
factors

Figure 1.2 The WHO International Classification of Functioning, Disability and Health (ICF)

As illustrated in Figure 1.2, the relationships between the core constructs and the relationships between each construct and the health condition is bidirectional. This means, for example, that impairment can affect activity limitation but also that activity limitation can affect impairment. For instance, a person with joint degeneration (impairment) may fail to exercise the joint (activity limitation) because movement is painful. As a consequence of inactivity, the muscles around the joint weaken (increasing impairment), which makes movement more difficult (increased activity limitation). In addition to the relationships between the core constructs and the health condition, the ICF states that contextual factors, namely ‘environmental factors’ and ‘personal factors’, further modify these relationships. These contextual factors provide the means by which the ICF can accommodate the role(s) for an individuals’ life situation and life experience in the process of disablement. Consequently, functioning and disability are viewed as a result of the dynamic interaction between the health condition, the individual and the environment.

1.3.1.1 <u>The ICF taxonomy</u>

The ICF has multidisciplinary appeal. It provides a standardised common language to communicate health and health-related issues internationally across varying disciplines. The taxonomy details what needs to be measured to operationalise the ICF. That is to say the ICF describes in some detail what bodily functions and structures should be measured to index impairment, and what activities and social participatory situations should be measured to index activity limitation and participation restriction. Similar details are available for the environmental component of the contextual factors construct; however the content of the personal factors component is yet to be detailed. These measurement domains are termed ‘categories’ within the ICF terminology and the ICF taxonomy currently comprises 1545 categories, each of which define a necessary aspect of functioning and disability (Boonen, Stucki, Maksymowych et al., 2009). In order to facilitate application of the ICF taxonomy, a list of the measurement categories that operationalise the ICF for a particular condition is compiled and forms the ICF core measurement set for that condition (Rat, Guillemin, & Pouchot, 2008).

1.3.1.2 <u>ICF core measurement sets</u>

ICF core measurement sets provide an internationally agreed list of ICF categories that should be considered and measured in clinical research and practice, and in multidisciplinary assessments of patients with a specific health condition (Cieza, Ewert, Ustun et al., 2004). ICF core sets have been developed for many chronic conditions including obesity (Stucki, Daansen, Fuessl et al., 2004), rheumatoid arthritis (Stucki, Cieza, Geyh et al., 2004) and osteoarthritis (Dreinhofer, Stucki, Ewert et al., 2004). The

ICF core measurement set for osteoarthritis (Dreinhofer et al., 2004) lists 55 ICF categories that are said to represent the typical spectrum of problems relevant to the functioning of people with osteoarthritis, including ‘sensation of pain’, ‘musculoskeletal structures related to movement’, ‘walking’ and ‘doing the housework’. These categories are in accord with the commonly reported symptoms of osteoarthritis identified in the literature (see Section 1.4.1). However, an ICF core set serves to inform *what* has to be measured in relation to the specific health condition but does not inform *how* to measure it. It has recently been recommended that the ICF categories included in a core set are used as the starting point to select items to be included in new measurement instruments for that health condition (Boonen et al., 2009). However, to enable existing health outcome measures to operationalise the ICF, linking rules have been developed to map the content of items from existing measures to the ICF categories (Cieza, Brockow, Ewert et al., 2002; Cieza, Geyh, Chatterji et al., 2005).

1.3.1.3 Linking items from health status measures to the ICF: the problem of discriminant validity

Application of the linking rules to map measurement items from existing health outcome measures to the ICF (Brockow, Cieza, Kuhlow et al., 2004; Rat et al., 2008; Weigl, Cieza, Harder et al., 2003) has highlighted a potentially serious problem for the framework. The ICF framework, presented in Figure 1.2, clearly differentiates between the constructs of activity and participation, providing corresponding theoretical definitions for each; however, the measurement categories used to index activity limitation are not distinct from the measurement categories used to index participation restriction. This has resulted in core measurement sets that list categories for the discriminant measurement of bodily function & structure (impairment) but a single component of activity & participation, combining the measurement of activity limitation together with participation restriction. The lack of discriminant measures for each of the theoretical constructs means that the core sets cannot be used to operationalise the ICF framework in a manner that enables model testing.

Pollard, Johnston & Dieppe (2006) emphasise this measurement concern in relation to testing the effect of a treatment. They argue that in order to identify all the potential effects on each of the ICF constructs, namely impairment, activity limitation and participation restriction, it is important to ensure that the outcome measure can actually measure each of the ICF constructs independently, i.e., without measurement confound. Therefore, using the example of joint replacement surgery for osteoarthritis, in order to identify the true effect of surgery on each health outcome identified by the ICF namely impairment, activity limitation and participation restriction, items with discriminant validity should be chosen. Consider the item about *pain on walking*. Within the ICF, pain has been shown to be an indicator of impairment (Cieza et al., 2004; Dreinhofer et al., 2004; Pollard, Johnston, & Dieppe, 2006), therefore, the true effect of joint replacement surgery on the impairment construct (i.e., pain) may be masked because of the item's concurrent measurement of the activity limitation construct (i.e., walking). Consequently, this item would not be a *pure* measure of impairment or activity limitation and thus not an appropriate item to detect independent treatment effects on each health outcome.

These empirical concerns about the ability to measure each ICF construct with discriminant validity have fuelled the debate regarding the conceptual and theoretical distinction of the activity and participation components of the ICF (Jette, Tao, & Haley, 2007; Jette, Haley, & Kooyoomjian, 2003; Perenboom & Chorus, 2003). That said, an alternative method of establishing the content validity of measurement items, namely Discriminant Content Validation, has identified conceptually pure measures of impairment, activity limitation and participation restriction (Dixon, Pollard, & Johnston, 2007; Dixon, Johnston, McQueen et al., 2008a; Pollard et al., 2006). However, the issue of whether activity and participation should be distinct constructs or whether they should be combined into a single construct remains a topic of active discussion.

1.3.2 Psychological Models: The Theory of Planned Behaviour

The re-conceptualisation of disability as behaviour brought disability into the sphere of health psychology. Before disability was conceptualised as behaviour, the psychological approach to disability was typically limited to the treatment of the emotional consequences of disablement. However, the concept of disability as behaviour enabled models of human behaviour and behaviour change to be applied to further our understanding of the factors that influence disability. The Theory of Planned Behaviour, developed within social psychology as a general model of human behaviour, has been extensively used to understand the factors that predict many types of health behaviour. As such it can be used to understand disability behaviour (Johnston, 1996; Johnston, 1997; Johnston, Bonetti, & Pollard, 2002).

1.3.2.1 The Theory of Planned Behaviour (TPB) model

The TPB (Ajzen, 1991) is one of the most extensively applied social cognition models in the study of health behaviours (Armitage & Conner, 2001; Godin & Kok, 1996; Hagger, Chatzisarantis, & Biddle, 2002; Sheeran, 2002). The theory proposes that behavioural intention and perceived behavioural control are the most proximal predictors of behaviour. Intentions represent an individual’s decision and motivation to perform a behaviour and indicate the conscious exertion of effort to act on that decision

(Ajzen, 1991). Perceived behavioural control reflects the perceived amount of control that an individual has over the behaviour, and is comparable with Bandura's notion of self-efficacy (Bandura, 1977). Ajzen claims that perceived behavioural control can have a direct and indirect influence on the performance of behaviour; stronger perceived behavioural control directly predicts more engagement in behaviour and, stronger perceived behavioural control predicts stronger intentions, which in turn predict more engagement in behaviour. Intention is also influenced by an individual's attitude towards the behaviour and their perception of whether or not other people think they should perform the behaviour (subjective norm) Figure 1.3 presents the full TPB model; the proximal predictors of behaviour are shown in bold typeface.

Attitude

Subjective Norm

Intention

Behaviour

Perceived Behavioural Control
(Self-Efficacy)

Figure 1.3 The Theory of Planned Behaviour adapted from Ajzen (2006)

1.3.2.2 Applications of the TPB to health-related behaviours

There is substantial evidence in support of the theory in the prediction of intention and behaviour. A large meta-analytic review of the TPB applied to a variety of behaviours, including health-related behaviours, found that the TPB accounted for 27% and 39% of the variance in behaviour and intention, respectively (Armitage & Conner, 2001). Further, reviews of the application of the TPB to behaviours of immediate relevance to disability, for example physical activity (PA) behaviours central to mobility disability, have also identified strong associations between the TPB variables and PA (Godin & Kok, 1996; Hagger et al., 2002).

While there is compelling evidence for the predictive value of the TPB, there is less evidence addressing the causal structure of the theory. For example, Hardeman et al (2002) reviewed the literature to identify TPB-based interventions to change health behaviours. This review could only identify 24 interventions, only 13 of which reported behavioural outcomes and only 12 of which actually used the TPB to develop the intervention. However, more recent work has shown the TPB can be used to develop interventions to change health-related behaviour. For example, a randomised controlled trial of a TPB-based intervention to increase healthy eating and PA in older adults employed a healthy living booklet designed to target perceived behavioural control, intention and promote goal setting (Kelley & Abraham, 2004); the intervention group made higher gains in intention and perceived behavioural control and in self-assessed healthy eating and PA compared to the control group (see Chapter 5, Section 5.1.1 for a detailed discussion of the TPB intervention literature).

1.3.3 Integrating the ICF and the TPB

In 1996, Johnston proposed integrating psychological theory, namely the TPB, into the original WHO-ICIDH model to advance understanding of the relationship between impairment and disability (Johnston, 1996). Following the development of the revised WHO-ICF model, the TPB was integrated to advance understanding of the relationship between impairment and activity limitation (Johnston et al., 2002) (see Figure 1.4).

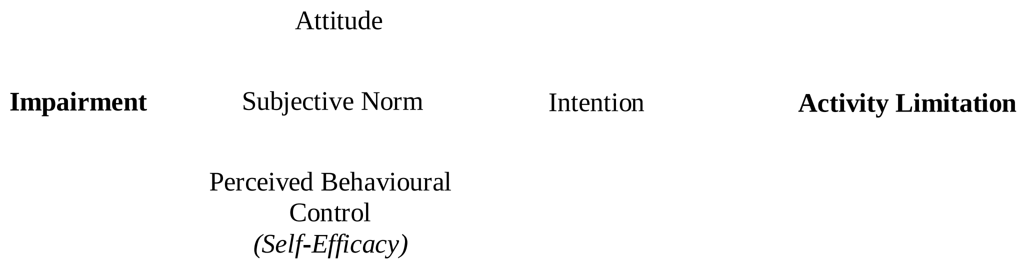


Figure 1.4 The integrated ICF/TPB model of disability

The measurement categories for the body function & structure (impairment) component and the combined activity & participation component of the ICF are defined in great detail. However, the measurement categories for the personal factors components of the contextual factors construct, defined as contextual factors that relate to the individual (World Health Organization, 2001), are yet to be agreed. A recent review of the ICF literature emphasised that confusion exists in the coding of individual attributes; for example, optimism, confidence and motivation are currently coded as part of ‘body functions & structure’ under the category of *temperament and personality*, whereas they may be better coded as ‘personal factors’ (Jelsma, 2009). In Johnston’s integrated model, such psychological constructs are construed as personal factors related to the individual. Intention (motivation) and control beliefs taken from the TPB are integrated into the ICF to further explain and predict the relationship between impairment and performance of an activity (Johnston et al., 2002). Consider a person with activity limitations in the form of mobility disability associated with degeneration of the hip joint, for example, the integrated ICF/TPB model indicates that the extent of their mobility disability is a function of the level of impairment, the strength of their motivation (intention) to walk and how much control they perceive they have over walking. To date, application of the integrated ICF/TPB model has shown that perceptions of control can predict activity limitations in patients with stroke (Bonetti &

Johnston, 2008; Johnston, Pollard, Morrison et al., 2004; Johnston, Bonetti, Joice et al., 2007), osteoarthritis (Dixon, Johnston, Rowley et al., 2008b) and chronic idiopathic axonal polyneuropathy (Schroder, Johnston, Teunissen et al., 2007).

1.4 Osteoarthritis and Disability

Many chronic health conditions are associated with disabilities of various forms. It can be argued that mobility disability is of particular concern because, in addition to the health and fitness benefits associated with walking (see Chapter 5 Section 5.1.4 for a review of the PA literature), the behaviour of walking is fundamental to being able to take part in many other activities. For example, being able to walk from the lounge to the bathroom is required if an individual is to perform other important activities of daily living such as toileting and bathing. The same behaviour expressed over longer distances, such as walking to and from a bus stop, is an important prerequisite for many social participatory activities, such as visiting the cinema. Osteoarthritis is a prevalent condition, associated with aging, one of the primary symptoms of which is mobility disability. As such, osteoarthritis is a suitable exemplar condition to study the factors that influence mobility disability within the integrated ICF/TPB framework. In addition, the availability of surgical interventions, which radically alter the impairment status in osteoarthritis, presents a somewhat unique opportunity to test any theoretical model of disability associated with a chronic illness. The large majority of chronic conditions do not afford the opportunity to study disability under conditions where the impairment is radically reduced or removed in this way. Osteoarthritis, therefore, is used as an exemplar health condition to examine the utility of the ICF and integrated ICF/TPB models of disability.

1.4.1 Osteoarthritis Symptoms and Consequences

Osteoarthritis is an incurable, long-term musculoskeletal condition. It is characterised by the disintegration of articular cartilage and the formation of new bone (Dekker, Boot, van der Woude et al., 1992). Knees and hips are the most commonly affected joints. Osteoarthritis is the most prevalent form of arthritis; 8.5 million people in the United Kingdom (UK) (Arthritis Care, 2004) and over 151 million people worldwide (World Health Organization, 2004) are estimated to live with osteoarthritis. The prevalence of osteoarthritis increases with age. At the age of 30 years, less than 5% of people show osteoarthritic changes as determined by radiography; by 65 years, at least 50% of people show such changes (Dekker et al., 1992).

The primary symptom of osteoarthritis is pain. Pain is typically worse on movement but can also be experienced at rest. It has been found to change and worsen over time in terms of its intensity and type, becoming unpredictable and resulting in significant avoidance of social and recreational activities (Hawker, Stewart, French et al., 2008). Qualitative studies have indicated that patients with osteoarthritis closely associate pain with movement and activity (Gooberman-Hill, Woolhead, MacKichan et al., 2007), although pain and activity can be distinguished in quantitative studies of patients with osteoarthritis awaiting joint replacement surgery (Dixon et al., 2008b). Activity limitation or disability is another important symptom of osteoarthritis. Findings from a large survey of people with osteoarthritis in the UK concluded that 81% experience constant pain or are limited in their scope to perform everyday tasks (Arthritis Care, 2004). Together these symptoms have a significant negative impact on an individual; reducing quality of life, limiting the performance of social roles, such as work or leisure pursuits (Dekker et al., 1992), and increasing the likelihood of further morbidity and mortality (Jordan, Arden, Doherty et al., 2003).

The disabling nature of osteoarthritis is highlighted in figures on the global burden of disease; in 2004, the WHO identified osteoarthritis as the sixth most common cause of disability globally (World Health Organization, 2004). Further, a United States based study examining the effects of reducing the prevalence of key fatal and nonfatal conditions, including coronary artery disease, stroke, cancer, diabetes and arthritis, projected that the greatest reduction in the number of adults with activity limitations would be achieved by decreasing the prevalence of arthritis, including osteoarthritis (Boult et al., 1996). In the absence of any interventions to reduce the prevalence of osteoarthritis, work to further our understanding of the factors that predict disability associated with osteoarthritis is of particular importance.

1.4.1.1 Management of osteoarthritis

Since osteoarthritis is an incurable condition, treatments are directed towards managing the symptoms and preventing their further deterioration (Dekker et al., 1992). The National Institute for Health and Clinical Excellence (National Institute for Health and Clinical Excellence, 2008) guidelines for the care and management of osteoarthritis recognise PA as a core treatment, and 57% of a surveyed UK sample reported using PA to help manage their condition (Arthritis Care, 2004). Arthritis patient education and self-management programmes have been widely evaluated and have been shown to reduce pain, depression, disability and healthcare utilisation (Lorig, Mazonson, & Holman, 1993; Lorig, Selenznick, Lubeck et al., 1989; Lorig, Sobel, Stewart et al., 1999; Lorig, Ritter, Laurent et al., 2004). A meta-analytic review of psychosocial interventions for arthritis identified positive effects on measures of coping, anxiety, joint swelling, depression, disability and pain self-efficacy (Dixon, Keefe, Scipio et al., 2007). Pharmacological management of osteoarthritis is also recognised; however, when patients become unresponsive to conservative analgesic treatments, total joint replacement surgery is an effective intervention for moderate to severe osteoarthritis of the hip or knee (Jordan et al., 2003; Zhang, Doherty, Arden et al., 2005). Total joint replacement surgery presents the opportunity to study change in mobility disability under conditions in which impairment has effectively been reversed.

1.5 Theory-Based Interventions

1.5.1 Interventions on Impairment to Reduce Disability: Total Joint Replacement

Total joint replacement (TJR) is an irreversible surgical intervention to treat advanced stages of osteoarthritis of the hip and knee in patients unresponsive to conservative treatments such as analgesics and exercise (Jordan et al., 2003; Zhang et al., 2005). TJR involves the removal of damaged tissue from the joint and replacement with an artificial joint made of metal and plastic. It is a frequently performed procedure; in England and Wales in 2006/2007 in excess of 58,000 primary total hip replacements (THR) and 62,000 primary total knee replacements (TKR) were recorded (National Joint Registry, 2007).

TJR is an effective procedure resulting in significant improvements in health-related quality of life (Ethgen, Bruyere, Richy et al., 2004; March, Cross, Lapsley et al., 1999; Shields, Enloe, & Leo, 1999). Moreover, large scale reviews indicate that up to 84% of patients report being pain free after THR and 89% of patients report good or excellent outcomes up to five years after TKR (Jordan et al., 2003; Zhang et al., 2005). Despite the general effectiveness of TJR, these figures indicate that some patients do not experience such improvements and that variability in the success of surgery exists. This variability has led to investigation of the factors that may account for individual differences in outcome and recovery from TJR, such as surgical and nonsurgical clinical factors, demographic factors, and psychological factors including specific beliefs and expectations (Fortin, Clarke, Joseph et al., 1999; Jones, Beaupre, Johnston et al., 2007; Kopp, Bonatti, Haller et al., 2003; Orbell, Johnston, Rowley et al., 1998). The ICF is a theoretical model of health outcomes that recognises the possible influence of clinical and psychological factors on health outcomes; therefore, it is an appropriate framework to investigate the relationship between cognitive beliefs such as patient expectations of TJR and recovery after surgery. Furthermore, as TJR is generally an effective intervention to reduce disability in people with osteoarthritis, understanding the factors that predict the provision of such surgery is important for any programme of work that aims to understand disability associated with osteoarthritis.

1.5.1.1 Indications for TJR

At present, there is little consensus over the indications for THR (Dreinhofer, Dieppe, Sturmer et al., 2006) or TKR (Dieppe, Basler, Chard et al., 1999). Dreinhofer et al (Dreinhofer et al., 2006) concluded that, in general, pain and functional limitations are the most important indicators, however other indicators have also been identified in the literature such as the presence of some radiographic change (Jordan et al., 2003; National Insitute of Health Consensus Panel, 1995; Zhang et al., 2005), patient motivation (Mancuso, Ranawat, Esdaile et al., 1996) and comorbidity (Maillefert, Roy, Cadet et al., 2008). In the absence of internationally agreed objective indication criteria for TJR, health professionals are likely to apply their own personal clinical decision making process to ascertain whether a patient qualifies for surgery. This individual decision making is likely to be influenced by the clinician’s idiosyncratic model of osteoarthritis, its symptoms and the likely effectiveness of surgery for a given patient.

Whilst many studies have identified factors that influence provision of TJR surgery, this work lacks a coherent theoretical framework. It is possible that the integrated ICF/TPB model of disability or its constituent parts (ICF & TPB) may provide a useful framework for understanding the factors that predict TJR, and thus further our understanding of the process of disablement, treatment and recovery in osteoarthritis. Indeed, even though much of the evidence base in this area is atheroetical, it has revealed that factors such as social participation play a role in clinical decision making around TJR. Factors such as social participation are possibly consistent with the concept of participation restriction within the ICF and integrated ICF/TPB models. For example, referring physicians appear to place more importance on indicators relating to quality of life and social involvement than orthopaedic surgeons, whilst orthopaedic surgeons placed more importance on indicators relating to the extent of joint damage (Dreinhofer et al., 2006). Similarly, mobility disability is also an important criteria for referral by family physicians for consultation with a rheumatologist; family physicians but not rheumatologists agreed that they were more likely to refer patients for TKR if walking was limited to less than one block without pain (Coyte, Hawker, Croxford et al., 1996). These findings suggest that disability-related (activity limitation and participation restriction) indicators are more important for family physicians, whereas impairment-related indicators are more important for consultants. This difference may reflect the typical focus of consultations between patients with osteoarthritis and each type of health professional. For example, consultations with general physicians may be more likely to include factors relating to the daily management of osteoarthritis, compared to consultations with orthopaedic consultants whose focus and experience may be more dominated by the impairment and pending surgical procedure.

1.5.1.2 Prioritisation for TJR

The current demand for TJR is high and with an aging population this demand is increasing (Birrell, Johnell, & Silman, 1999). Even though THR and TKR are frequently performed procedures, in countries with publicly funded healthcare systems and concurrent limited resources such as the UK, the provision of TJR does not always meet demand and waiting times exist. For example, the Musculoskeletal Services Framework in England (Department of Health, 2006) has developed the “18-week patient pathway” that states that patients will not wait more than 18 weeks from referral to hospital treatment. However, patients have to be prioritised on those waiting lists.

The evidence for the effect of the length of waiting time for TJR on health-related quality of life is mixed. Mahon, Bourne & Rorabeck’s (2002) prospective study of patients awaiting THR found that patients receiving surgery within 6 months after referral realised greater gains in health-related quality of life and mobility than patients waiting more than 6 months, suggesting that the priority assigned to patients is an important determinant of quality of life. In contrast, a randomised controlled trial of THR patients (Tuominen, Sintonen, Hirvonen et al., 2009) concluded that patients in the short waiting time group did not differ to patients in the longer waiting time group on health and quality of life outcomes measured 3 and 12 months after surgery. However, those in the short waiting time group did reach better quality of life 3 months earlier. This finding suggests that although a longer waiting time does not necessarily limit the potential gains in quality of life following THR, earlier surgery clearly provides the opportunity for those gains to be achieved sooner; thus, prioritisation is an important determinant of patient quality of life. Unfortunately, in addition to a lack of internationally agreed objective indications for TJR, there is a lack of internationally accepted evidence-based guidelines to aid clinicians when assigning priority to patients for surgery. Applying an international theoretical framework, such as the ICF, to investigate the factors that influence prioritisation for TJR is required.

1.5.2 Interventions on Cognitions to Reduce Disability and Increase PA

The integrated ICF/TPB model identifies cognitions as important direct predictors of disability and as important mediators of the effect of impairment on disability. Thus, the integrated model indicates that disability could potentially be reduced through interventions that target cognitions in relation to disability behaviour, without the need for concomitant reductions in impairment. Indeed, experimental studies have shown that simple interventions to increase control cognitions can reduce activity limitations in people with chronic low back pain (Fisher & Johnston, 1996). The cognitions to be targeted by an intervention would be the proximal predictors of behaviour within the TPB, namely intention and perceived behavioural control, the latter predictor measured as either perceived controllability or self-efficacy beliefs (Ajzen, 2002). In people with mobility problems, such as those associated with osteoarthritis, an intervention to increase these TPB cognitions in relation to disability behaviours would be expected to reduce disability and increase PA behaviours. In comparison, in healthy people without mobility problems or any impairment, an intervention to increase these TPB cognitions in relation to PA behaviours would be expected to increase PA behaviours.

1.5.2.1 Limitations of group design studies

At present, the majority of intervention studies in health psychology are based on group designs. Group-based designs have successfully identified differences between individuals, however group-based designs cannot be used to identify within-individual intervention effects (Molenaar, 2004). For example, a typical group-based study applying the TPB to predict PA examines whether stronger intentions and higher perceived behavioural control predict more PA. Furthermore, a typical group-based experimental study applying a TPB-based intervention to increase PA examines whether an increase in the targeted TPB construct and subsequent PA is identified in the intervention group. In both predictive and experimental studies, data from all participants in each group is pooled for statistical analyses and the effect of the intervention is evaluated by comparing the average effects across groups. Therefore, findings relate to the group and cannot be extrapolated to represent the specific effect on any individual within the group (Ottenbacher, 1990). Moreover, the utility of a model or theory to predict and explain behavior change within an individual is not examined.

An alternative to the group-based design is the study of individuals with an *n-of-1* (single subject) design, as suggested in the Medical Research Council (MRC) Complex Interventions Guidance (Craig, Dieppe, Macintyre et al., 2008). In contrast to the group design examining *between*-individual difference, the *n-of-1* design examines *within*-individual difference and has the potential to test theory and the effectiveness of an intervention within an individual. This may be of particular use in complex interventions typically used in the management of chronic conditions, which may require an intervention to be adapted to the needs and capabilities of each individual patient. A demonstration that the theoretical framework(s) used in group-based studies have predictive validity in relation to individual behaviour change, including disability and physical activity behaviour, would be of particular interest.

1.6 Aims of This Thesis

This thesis employs the theoretical frameworks of the ICF, the TPB and the integrated ICF/TPB model to investigate disability and PA behaviours. Thus, an overarching aim of the thesis is to facilitate the development of a cumulative evidence base by the application of these theoretical frameworks. General aims that relate to particular studies reported in this thesis are:

1. Can the ICF be used to understand clinical prioritisation tools for TJR, which were developed in the absence of any explicit theoretical framework (Chapter 2)?

2. Can the ICF and a simplified version of the integrated ICF/TPB model be used to understand clinical decision making in relation to the prioritisation for THR (Chapter 3)?

3. Can the ICF, as a model of health outcomes, be used to classify pre-operative patient expectations of outcome after THR and to investigate the relationship between patient expectations and recovery after surgery (Chapter 4)?

4. Is the TPB predictive of PA *within* individuals with mobility problems due to osteoarthritis (Chapter 6) and *within* healthy individuals without mobility problems (Chapter 7)?

5. Is there support for the candidate causal pathways in the TPB *within* individuals with mobility problems due to osteoarthritis (Chapter 6) and *within* healthy individuals without mobility problems (Chapter 7)?

Chapter 2 - Prioritisation for total joint replacement – operationalisation of clinical priority assessment tools

2.1 Abstract

Background: Total joint replacement surgery is an effective procedure producing measurable reductions in pain and disability. However, provision does not always meet demand, therefore, waiting times exist and patients have to be assigned priority for surgery. Prioritisation scoring systems have been developed to aid clinicians with the prioritisation decision. This study examined whether two prioritisation tools are compatible with the International Classification of Functioning, Disability and Health framework (ICF).

Method: Thirteen health professionals judged the 20 items comprising the New Zealand Priority Criteria Project tool and 11 health professionals judged the 25 items comprising the Western Canada Waiting List Project tool. The method of discriminant content validation was employed which requires participants to judge whether each item measures each of the ICF constructs, impairment, activity limitation and participation restriction, and to rate their confidence in these judgements. Judgements were weighted and one sample *t*-tests were employed to determine which ICF constructs were being measured by each item.

Results: Both prioritisation tools contained items that were pure measures of impairment, activity limitation and participation restriction. However, judges showed a lack of agreement about whether an item measured any of the ICF constructs in relation to a total of 18 items (9 from each tool); these items were deemed unclassifiable in relation to the ICF.

Conclusions: One of the aims of the prioritisation tools is to reduce variability in clinical decision making. The lack of agreement between judges in relation to 18 items suggests these items are unlikely to facilitate consistency in clinical decision making. It is suggested that the use of complex and lengthy text within items underlies this lack of agreement between judges.

2.2 Introduction

Total joint replacement (TJR) is a frequently performed and effective treatment for osteoarthritis of the knee and hip. However, the demand for TJR is high; therefore, the provision of TJR within publicly funded healthcare systems with limited resources does not always meet this demand. As a result, patients are placed on waiting lists for surgery such as those that exist within the National Health Service in the UK. Like many other public healthcare systems in other countries, the number of people on a waiting list for surgery plus the length of time they have been waiting has been of increasing interest within the UK government (Health Policy and Economic Research Unit, 1998). Guidelines and targets are set for maximum waiting times such as the “18-week patient pathway” in the Musculoskeletal Services Framework in England (Department of Health, 2006), whereby patients wait no more than 18 weeks from GP referral to hospital treatment. These guidelines are designed to reduce rather than abolish waiting times therefore waiting still exists. Having to wait for surgery means that the opportunity to realise the potential gains in health-related quality of life associated with TJR is delayed (Tuominen et al., 2009); therefore, the process of assigning priority and a corresponding position on the waiting list is an important factor in determining patient quality of life.

At present, there is a lack of consensus over the indications for total knee replacement (TKR) (Dieppe et al., 1999) and total hip replacement (THR) (Dreinhofer et al., 2006). Furthermore, no internationally accepted evidence-based guidelines exist to aid clinicians in the prioritisation decision. For example, the British Medical Association (Fricker, 1999) proposes the use of five categories ranging from emergencies to low priorities in order to manage waiting lists; however, at present, it does not propose an explicit or standardised procedure to guide this clinical decision. The lack of international standardisation and observed inconsistent judgements within and between clinicians may mean that patients with identical clinical characteristics have very different waiting periods, which has been suggested as one explanation for the wide variation observed in the provision of TJR within and between countries (Coyte et al., 1996; Naylor, Williams, & Ontario panel on hip and knee replacement, 1996).

In order to address the issue of standardising the prioritisation process, priority scoring systems for TJR and other elective surgeries have been developed in several countries (see MacCormick, Collecutt, & Parry, 2003 for a review). The main argument for the use of clinical priority scoring systems is to make the management of waiting lists transparent, utilising explicit criteria to assign priority, thus ensuring that clinical need and ability to benefit rather than crude waiting time are the basis for receiving surgery (Edwards, 1999; MacCormick et al., 2003; Health Policy and Economic Research Unit, 1998).

The New Zealand Priority Criteria Project (NZPC: Hadorn & Holmes, 1997) has been at the forefront of the production of clinical priority assessment tools to aid management of waiting lists for many elective surgeries including TJR. The severity of the patient's condition and subsequent clinical urgency is assessed by criteria which are summed to produce a score from 0 (i.e., lowest priority) to 100 (i.e., highest priority). The clinician is required to rate the patient on eight criteria namely*, degree of pain, occurrence of pain, time walked, other functional limitations, pain on examination, other abnormal findings, multiple joint disease* and *social factors*. For each criterion, the clinician chooses the item that best describes the patient's current health and functioning and the item's corresponding score is noted. Scores are summed across all criteria to determine the patient's overall priority score. In addition to its use to manage waiting lists for TJR in New Zealand, the NZPC tool for TJR has been internationally recognised as a valuable scoring system in research being used in studies of population requirement for THR (Frankel, Eachus, Pearson et al., 1999) and TKR (Juni, Dieppe, Donovan et al., 2003), waiting times (Harry, Nolan, Elender et al., 2000) and indications for THR (Maillefert et al., 2008).

In Western Canada a similar project group was established to address the problem of waiting list management. The Western Canada Waiting List Project (WCWL) developed a priority criteria scoring system for TJR based on the earlier work in New Zealand (Arnett, Hadorn, & Steering Committee of the Western Canada Waiting List Project, 2003). Like the NZPC tool, the WCWL tool produces a score from 0 to 100 to reflect the patient’s clinical urgency for TJR. Seven criteria are assessed, namely *pain on motion*, *pain at rest, pain on walking, other functional limitations, abnormal findings on physical examination, radiographic findings* and *social role*. Like the NZPC, the clinician is required to choose the item that best describes the patient’s current health and functioning for each criterion and the corresponding scores for each item are summed across all criteria to determine the patient’s overall priority score. The majority of research using the WCWL tool to date has been conducted by the WCWL steering group (Conner-Spady, Estey, Arnett et al., 2004; Conner-Spady, Estey, Arnett et al., 2005; de Coster, McMillan, Brant et al., 2007); however, evidence of its validity and reliability has begun to trigger its use and refinement internationally (Western Canada Waiting List Project, 2004).

The NZPC and the WCWL tools were developed through extensive clinical input and iterative exchange of clinical advisory groups (Derrett, Paul, Herbison et al., 2002; Noseworthy, McGurran, Hadorn et al., 2003); however neither was developed within a theoretical framework. In light of the current and likely future application of these prioritisation tools in international TJR research and practice, it would be useful to investigate how the tools fit into an international theoretical framework of health outcomes, namely the International Classification of Functioning, Disability and Health (ICF).

As discussed in Chapter 1 Section 1.3.1.2, an ICF core measurement set has been developed for osteoarthritis (Dreinhofer et al., 2004), helping inform the decision about *what* ICF categories to measure in osteoarthritis research and practice. The core set however, does not inform the decision about *how* to measure the identified ICF categories. ICF linking rules have been developed (Cieza et al., 2002; Cieza et al., 2005) and applied to identify the content of health status and outcome measures used in musculoskeletal research in relation to the ICF framework (Brockow et al., 2004; Rat et al., 2008; Weigl et al., 2003). The linking rules state that a single item from a measure can possess multiple concepts; thus, each concept within an item should be linked individually (Cieza et al., 2002). Despite the intuitive appeal and favourable evaluations of these linking rules, the possibility that a single item can potentially be mapped to ICF categories taken from different components of the ICF, means that there is a high risk of measurement confound within the instruments, as discussed by Pollard et al (2006) (see Chapter 1 Section 1.3.1.3). In order to identify all the potential effects on each ICF component, when testing the effect of a treatment, pure measures of each ICF construct are needed.

The methodology of Discriminant Content Validation (DCV) has been developed and applied to investigate the content of various orthopaedic and chronic pain health-status measures, specifically whether a single or multiple ICF constructs is measured by a given item (Dixon et al., 2007; Dixon et al., 2008a; Pollard et al., 2006). Expert judges are required to decide whether the item matches the theoretical definition of each of the

ICF constructs and then to rate their confidence about this decision. Conclusions can then be drawn regarding whether a measure contains *pure* items measuring a single ICF construct and/or *mixed* items measuring multiple ICF constructs.

This study uses the DCV methodology to explore which factors of patient health and functioning, as defined by the ICF framework, are measured by two clinical prioritisation tools for TJR, namely the NZPC and the WCWL. The objective is to identify the content of the items within each tool in relation to the ICF. Knowledge relating to the relative number of items in the prioritisation tools mapped to each of the ICF constructs would provide an indication of the importance given to each component of health functioning when assigning priority to a patient for TJR.

2.2.1 Research Questions

1. Can the New Zealand Priority Criteria Project (NZPC) tool measure each of the ICF constructs of impairment (I), activity limitation (A) and participation restriction (P) with discriminant validity?

2. Can the Western Canada Waiting List Project (WCWL) tool measure each of the ICF constructs of impairment (I), activity limitation (A) and participation restriction (P) with discriminant validity?

2.3 Methodology

2.3.1 Design

Participants matched each of the 20 items from the NZPC and each of the 25 items from the WCWL to the theoretical definition of the ICF constructs of I, A and P.

2.3.2 Participants

A convenience sample of 13 health professionals (academics in health psychology, clinical psychology and health services research) from the University of Stirling and the University of Aberdeen were invited to take part in the study. Thirteen participants accepted the invitation and completed the task for the NZPC. Eleven from this sample also completed the task for the WCWL. The number of participants falls within the recommendation of between 3 and 20 expert judges for content judgement tasks (McGartland Rubio, Berg-Weger, Tebb et al., 2003). Three participants had no prior knowledge of the ICF model; the remaining had some or extensive knowledge.

2.3.3 Materials

The definition of the ICF constructs I, A and P, as provided by the WHO (World Health Organization, 2001) are shown in Table 2.1. The items compiling the NZPC tool and the WCWL tool can be found in Appendices 1 and 2, respectively.

Table 2.1 Definitions of the WHO-ICF constructs

ICF Construct	Definition
Impairment (I)	Problems in body function or structures such as a significant deviation or loss
Activity limitation (A)	Difficulties an individual may have in executing activities
Participation restriction (P)	Problems an individual may experience in involvement in life situations

2.3.4 Procedure

Participants recruited from the University of Stirling performed the task as part of a monthly Health Psychology research group meeting. They were presented with an overview of the ICF model and informed of the purpose and application of the NZPC and WCWL as clinical prioritisation tools for TJR. Participants recruited from the University of Aberdeen were judges known to have knowledge of and experience using the ICF and the DCV methodology and hence, did not receive this briefing.

All participants were provided with two questionnaires; one detailing the NZPC items and the other detailing the WCWL items. Both questionnaires presented the definitions of I, A and P at the top of each page for easy reference. Items were listed in the same order as found in the tools. Participants were asked to read and consider each item carefully and to decide if each item matched each of the definitions of I, A and P, indicating their responses by circling either YES or NO. They were then asked to rate their confidence for each of these YES/NO judgements on a scale ranging from 0% i.e., not at all confident to 100% i.e., completely confident. The scale was in increments of 10%. Consequently, each participant made three judgements about each item: (1) does the item match the impairment definition and how confident am I about this judgement? (2) does the item match the activity limitation definition and how confident am I? and; (3) does the item match the participation restriction definition and how confident am I?

2.3.5 Analyses

Each judgement was weighted by first coding it as +1 for a match and -1 for no match. This was then multiplied by the corresponding confidence rating expressed as a proportion. For example, an item judged as a match to A with 80% confidence was assigned a value of .80; whilst an item judged as no match to I with 50% confidence was assigned a value of -.50. Thus all weighted judgements fell between -1 and +1. Missing values were coded as 0. Using one-sample *t*-tests, each item was classified as related to each possible construct when its weighted value was found to be significantly greater than 0 and each item was classified as *not* related to a construct when its weighted value was significantly less than 0. Subsequently, each item was classified to one of the seven possible construct combinations: I, A, P, impairment & activity limitation (IA), activity limitation & participation restriction (AP), impairment & participation restriction (IP) or impairment & activity limitation & participation restriction (IAP). Hochberg's correction was applied to correct for multiple tests (Hochberg, 1988).

Intraclass correlation coefficients (ICC) using a 2-way mixed model with measure of consistency were calculated to examine the agreement in judgements between the judges. The ICC ranges from 0 for complete independence of observations (i.e., complete *disagreement* between judges) to 1 for complete dependence of observations (i.e., compete *agreement* between judges)(Cohen, Cohen, West et al., 2003). The ICCs were calculated using the weighted values with 95% confidence intervals (CI) and compared against Cohen's (Cohen, 1992) large effect size criterion of .80 using an *F* test, as suggested by McGraw and Wong (McGraw & Wong, 1996). Inter-rater reliability was assessed for all items on each measure, i.e., 20 items for the NZPC and 25 items for the WCWL, and for judgements related to each construct i.e., consistency between judges for I judgements, A judgements and P judgements separately.

2.3.6 Ethics

Ethical approval was obtained from the Department of Psychology, University of Stirling.

2.4 Results

2.4.1 New Zealand Priority Criteria Project (NZPC) Tool

2.4.1.1 Inter-rater reliability

Inter-rater reliability for judgements across all items on the NZPC was high with an ICC of 0.88 (CI 0.84-0.92). The ICC for I judgements was 0.89 (CI 0.80-0.95), for A judgements was 0.89 (CI 0.81-0.95) and for P judgements was 0.89 (CI 0.81-0.95). All ICC values significantly exceeded the large effect size criterion of .80 (p =.001). The unique contribution of each judge on the ICC was examined by observing the extent of change in the value if the judge were removed from the analysis. All participants were found to be performing similarly and performance was not associated with their level of knowledge of the ICF model; hence all participants were included in the DCV analyses.

2.4.1.2 Discriminant content validation

After correction for multiple tests, 11 of the 20 items were classified to the ICF (see Table 2.2). Eight of them were judged to be a pure measure of the one of the ICF constructs, i.e., items 18 and 19 measured I, items 3, 4, 13, 14 and 15 measured A, and item 20 measured P. Three items were judged to be mixed: item 5 measured IA and items 12 and 16 measured AP. No items were judged to measure either IP or IAP.

There was no agreement between participants regarding what constructs were measured by the remaining 9 items however item 10 was judged to *not* measure either I or P. The unclassifiable items belonged to the following criteria: *degree of pain* (items 1 and 2), *occurrence of pain* (items 6, 7, 8 and 9), *time walked* (items 10 and 11) and *pain on examination* (item 17). Items 2, 8, 9 and 17 had been judged as measuring I but after correcting for multiple tests, they failed to reach significance.

Table 2.2 Classification of NZPC items to the ICF

Item	Classification	Impairment		Activity limitation		Participation restriction	
		t	*p*	*t*	*p*	*t*	*p*
1	None	3.14	*ns*	1.29	*ns*	0.54	*ns*
2	None	3.64	*ns*[a]	1.44	*ns*	1.37	*ns*
3	A	3.11	*ns*	20.14	0.001	2.77	*ns*
4	A	3.04	*ns*	17.48	0.001	1.94	*ns*
5	IA	4.87	0.001	21.35	0.001	3.19	*ns*
6	None	0.89	*ns*	0.20	*ns*	-2.94	*ns*
7	None	1.23	*ns*	1.80	*ns*	-2.00	*ns*
8	None	3.72	*ns*[a]	3.08	*ns*	1.05	*ns*
9	None	3.67	*ns*[a]	-2.74	*ns*	-1.32	*ns*
10	None[b]	-8.03	0.001	0.86	*ns*	-7.56	0.001
11	None	-2.17	*ns*	3.29	*ns*	1.85	*ns*
12	AP	-1.86	*ns*	14.30	0.001	4.09	0.001
13	A	-1.40	*ns*	14.17	0.001	2.34	*ns*
14	A	-0.22	*ns*	28.96	0.001	2.95	*ns*
15	A	0.84	*ns*	11.14	0.001	1.87	*ns*
16	AP	-0.80	*ns*	16.29	0.001	16.40	0.001
17	None	3.75	*ns*[a]	-0.08	*ns*	-1.56	*ns*
18	I	22.05	0.001	-0.89	*ns*	-1.93	*ns*
19	I	14.45	0.001	-2.95	*ns*	-2.95	*ns*
20	P	-1.15	*ns*	1.71	*ns*	15.80	0.001

Note. I = impairment; A = activity limitation; P = participation restriction; IA = impairment & activity limitation; AP = activity limitation & participation restriction; IP = impairment & participation restriction; IAP = impairment & activity limitation & participation restriction; None = item not classified to any construct; [a] item is not significant after Hochberg's correction for multiple tests; [b] item is significantly classified as *not* I or P.

2.4.2 Western Canada Waiting List Project (WCWL) Tool

2.4.2.1 <u>Inter-rater reliability</u>

Inter-rater reliability for judgements across all items on the WCWL was high with an ICC of 0.87 (CI 0.83-0.91). The ICC for I judgements was 0.87 (CI 0.78-0.93), for A judgements was 0.89 (CI 0.81-0.94) and for P judgements was 0.90 (CI 0.83-0.95). All ICC values significantly exceeded the large effect size criterion of .80 (p =.002). As with the NZPC tool, all participants performed similarly on their judgements of the

WCWL regardless of their level of knowledge of the ICF model and therefore, all participants were included in the DCV analyses.

2.4.2.2 Discriminant content validation

After correction for multiple tests, 16 of the 25 items were classified to the ICF (see Table 2.3). Fourteen items were judged to be a pure measure of one of the ICF constructs, i.e., items 17, 18, 19, 20, 21 and 22 measured I, items 3, 9, 10, 11, 13 and 14 measured A, and items 23 and 25 measured P. Two items were judged to be mixed: items 15 and 24 measured AP. No items were judged to measure IA or IAP. In addition to being judged as measuring I, item 19 was classified further as *not* measuring A or P and item 21 was classified further as *not* measuring P.

Participants did not agree on what constructs were measured by the remaining 9 items however item 16 was judged to *not* measure P. The unclassifiable items belonged to the following criteria: *pain on motion* (items 1 and 2), *pain at rest* (items 4, 5, 6 and 7), *pain on walking* (item 8), *other functional limitations* (item 12) and *abnormal findings on physical examination* (item 16). Five judgements failed to reach significance after correcting for multiple tests: items 6 and 7 had been judged as measuring I, item 2 as A, item 12 as *not* I, and item 20 as *not* A.

Table 2.3 Classification of WCWL items to the ICF

Item	Classification	Impairment		Activity limitation		Participation restriction	
		t	*p*	*t*	*p*	*t*	*p*
1	None	0.59	*ns*	1.34	*ns*	-3.08	*ns*
2	None	1.98	*ns*	3.85	*ns*[a]	-0.54	*ns*
3	A	2.67	*ns*	20.43	0.001	-0.11	*ns*
4	None	1.38	*ns*	-0.35	*ns*	-2.99	*ns*
5	None	2.48	*ns*	-0.50	*ns*	-2.46	*ns*
6	None	4.03	*ns*[a]	0.57	*ns*	-0.84	*ns*
7	None	4.05	*ns*[a]	0.96	*ns*	-0.58	*ns*
8	None	0.81	*ns*	1.62	*ns*	-1.98	*ns*
9	A	0.55	*ns*	15.55	0.001	-1.37	*ns*
10	A	1.66	*ns*	21.70	0.001	0.21	*ns*
11	A	-0.58	*ns*	14.83	0.001	-0.35	*ns*
12	None	-3.93	*ns*[a]	1.26	*ns*	0.91	*ns*
13	A	-1.51	*ns*	11.57	0.001	2.33	*ns*
14	A	-1.02	*ns*	13.43	0.001	2.45	*ns*
15	AP	-0.40	*ns*	23.35	0.001	15.78	0.001
16	None[b]	1.62	*ns*	-0.76	*ns*	-4.40	0.001
17	I	17.13	0.001	-0.54	*ns*	-1.64	*ns*
18	I	19.48	0.001	1.12	*ns*	-1.48	*ns*
19	I[c]	4.94	0.001	-4.86	0.001	-4.99	0.001
20	I	5.09	0.001	-4.03	*ns*[a]	-3.03	*ns*
21	I[b]	19.69	0.001	-2.21	*ns*	-6.58	0.001
22	I	22.75	0.001	-2.09	*ns*	-1.14	*ns*
23	P	-2.92	*ns*	2.14	*ns*	5.35	0.001
24	AP	-1.58	*ns*	4.49	0.001	4.08	0.001
25	P	-0.91	*ns*	2.92	*ns*	22.75	0.001

Note. I = impairment; A = activity limitation; P = participation restriction; IA = impairment & activity limitation; AP = activity limitation & participation restriction; IP = impairment & participation restriction; IAP = impairment & activity limitation & participation restriction; None = item not classified to any construct; [a] item is not significant after Hochberg's correction for multiple tests; [b] item is significantly classified as *not* P; [c] item is significantly classified as *not* A or P.

2.5 Discussion

This study found that both clinical priority scoring systems for TJR, namely the NZPC tool and the WCWL tool, contain *pure* items that can discriminantly measure each of the ICF constructs I, A and P. These *pure* items can therefore be used to reliably and validly measure a single ICF construct without contamination from concurrent measurement of another ICF construct. In addition, both tools were found to conflate measurement with *mixed* items judged as measuring multiple ICF constructs. This finding is consistent with previous studies; for example, 12 of 13 osteoarthritis-specific and general health measures were found to include mixed items (Pollard et al., 2006).

However, previous DCV analyses of existing health outcome measures only identified a few, if any, items that could not be classified to any of the ICF constructs (Dixon et al., 2007; Pollard et al., 2006; Dixon et al., 2008a). In contrast, the current study identified 18 items, within the two instruments, that could not be classified as either pure or mixed ICF items (45% of the NZPC items and 36% of the WCWL items), which indicates a lack of agreement between judges about whether the items measure any of the ICF constructs or a combination thereof. In addition to the lack of agreement about whether an item *does* measure a construct or combination thereof, the DCV methodology also allows for judges to agree that an item *does not* measure a construct or combination thereof. However, the current study found that judges only agreed on 2 of the 18 unclassifiable items in this way (i.e., item 10 of the NZPC *did not* measure I or P and item 16 of the WCWL *did not* measure P). Therefore, not only was there a lack of agreement between judges on what ICF constructs were being measured in 18 items, but there was also a lack of agreement on what ICF constructs were *not* being measured in 16 of these items.

Examination of the 18 unclassifiable items within the two instruments revealed that, according to the criteria specified within each instrument, the majority of the items were assessing pain. For example, of the 9 unclassifiable items identified in the NZPC tool, 4 items were within the criterion of *occurrence of pain* and 2 items within the criterion of *degree of pain*. Similarly, of the 9 unclassifiable items identified in the WCWL tool, 4 items were within the criterion of *pain at rest* and 2 items were within the criterion of *pain on motion*. Within the ICF, measurement items assessing the ‘sensation of pain’, such as those typically used in osteoarthritis-specific outcome measures, are used to measure the ICF construct of I (Brockow et al., 2004; Cieza et al., 2004; Dreinhofer et al., 2004). Furthermore, the method of DCV has consistently shown that items from pain measures assessing pain intensity are consistently assigned to the theoretical definition of I (Dixon et al., 2007; Pollard et al., 2006). Therefore, the current finding that these items were not judged to measure I (either as a pure measure or a mixed measure with another construct) is contrary to previous studies.

However, closer examination of the wording in these items revealed that the items clearly incorporate other concepts in addition to the sensation of pain, which may explain why these items were not classified as pure measures of I. For example, item 8 of the NZPC tool addressing *occurrence of pain* states “With all walking, mostly day pain” and item 5 of the WCWL tool addressing *pain at rest* states “Mild pain at rest: patient experiences some pain but it does not disturb their rest when they are sitting or lying down. Pain does not cause sleep disturbance”. The DCV methodology recognises the possibility that an item may measure multiple ICF constructs and for that reason each item is judged in relation to each of the ICF constructs. Therefore, the item “With all walking, mostly day pain” includes the concept of pain and the concept of walking, and therefore would perhaps have been expected to have been classified as a mixed item measuring I and A. It is possible that the inclusion of non-pain concepts in an item within the criterion of the instrument that claims to assess pain resulted in confusion in the judgement decision, which may have led to disagreement between participants and the items being deemed unclassifiable.

Furthermore, some items were classified as pure measures of constructs even though the concepts within the items measure multiple constructs. For example, item 2 of the NZPC states “Moderate pain: patient is active but has had to modify or give up some activities because of pain” and item 9 of the WCWL states “Patient can walk between 1 and 5 blocks but then must stop due to the pain”. Both of these items include the concepts of activities (walking) and pain, yet were classified as pure measures of A suggesting that even though pain was explicitly referred to in each item and is generally regarded as a measure of I (Pollard et al., 2006), the participants did not judge the items to measure I. The DCV methodology requires that the judge decides whether the item measures each of the ICF constructs independently and rates their confidence over this judgement, but perhaps, the decision making process is more complex than this. It is possible that participants’ decisions involved consideration of what concept they perceived to be most salient within the item. This may reflect poor explanation of the methodology to the participants by the researcher; however, this is unlikely as the procedure was explained in the same way as used in previous studies (Dixon et al., 2007; Dixon et al., 2008a; Pollard et al., 2006).

A possible explanation for the unexpected classification of some items and the large number of unclassifiable items, within the two instruments examined using the DCV methodology in the current study, may lie in the complexity and length of the text in the items in the prioritisation instruments. To date, DCV studies have examined measures with items that have been short and simple, many of which are used as outcome measures available to patient self-report. For example, the following two items are taken from the Chronic Pain Grade questionnaire, which has been examined using the DCV methodology (Dixon et al., 2007):

> “In the past 6 months, how intense was your worst pain rated on a 0-10 scale where 0 is ‘no pain’ and 10 is ‘pain as bad as it could be?”
>
> “About how many days in the last 6 months have you been kept from your usual activities (work, school or housework) because of this pain?”

In contrast, the items in the prioritisation tools are longer and more complex, and are designed to be used by clinicians. For example, the following two items are taken from the WCWL:

> “None/mild pain on motion: patient can move about walking and bending. They may experience some pain but it does not prevent any activity. They usually do not require pain medication.”
>
> “Moderate pain on motion: patient can move about including walking and bending. They experience pain most of the time which limits their activities to some degree. For example, patients experience trouble walking up and down stairs or may be uncomfortable standing for long periods of time. They occasionally need pain medication.”

The length and complexity of the prioritisation tool items may mean that a participant’s ability to recall all of the item or his/her inability to hold the full content of the item in working memory affects the information they use to classify the item. If participants were recalling different parts of the item whilst making the judgements, then it is possible that this may explain the lack of agreement in judgements between participants and the number of unclassifiable items. Similarly, the possibility that the length and complexity of the wording in the items in the prioritisation tools made the judgement process more difficult for the participants in the current study, may also hold for clinicians when employing the instrument to assign priority to a patient. If this is the case, then the complex item wording might also introduce variability in clinician judgments for the same reason. This raises questions about whether the prioritisation tools can achieve one of their key aims, which is to reduce inter-clinician variability in clinical decision making and standardise the prioritisation of patients for surgery (MacCormick et al., 2003). Thus, it would be of interest to replicate the DCV study with clinicians as judges to see if the lack of consensus is evident in their matching of items from the prioritisation tools to the ICF constructs. Alternatively, it is always possible that clinicians, especially those working in the same healthcare team, may share a common interpretation of the items in the tools.

2.5.1 Summary and Implications

Both the NZPC and the WCWL prioritisation tools contain pure items measuring each of the ICF constructs, namely I, A and P. However, there were a large number of unclassifiable items in both of the tools, i.e., items about which there was variability in participants' judgements. The findings suggest that complex wording may introduce the potential for variability in response to an item and this is problematic for prioritisation tools which aim to standardise clinical decision making. An instrument designed to standardise clinicians' prioritisation decision should comprise items that are not open to individual interpretation or lengthy and complex because this may limit the individuals' capacity to hold the entire item's content in working memory.

The presence of many unclassifiable items precludes analysis of whether prioritisation for TJR is being made on the basis of I, A or P, or a combination thereof. Thus, it may be useful to employ other instruments in conjunction with these tools to enable investigation of the health outcomes used in the prioritisation decision. For example, during the consultation, the WOMAC, an osteoarthritis-specific health outcome measure which has already been mapped to the ICF framework (Pollard et al., 2006), could be employed in conjunction with the prioritisation tool. This would allow investigation of the relationship between a patient's assigned priority for TJR and the severity of osteoarthritis measured in terms of scores on pure I, A and P items of the WOMAC.

Chapter 3 - Prioritisation for total hip replacement – an exploratory investigation of the factors that influence prioritisation

3.1 Abstract

Background: Total hip replacement surgery (THR) is associated with significant reductions in pain and disability, therefore, prioritisation for this elective surgery is an important determinant of quality of life. This study employed the International Classification of Functioning, Disability and Health (ICF) model and the integrated ICF/Theory of Planned Behaviour (TPB) model to explore the factors that influence prioritisation for THR.

Method: Forty-two orthopaedic surgeons ranked two sets of patient vignettes in order of priority for THR. Set 1 comprised 8 vignettes that varied systematically by severity of the ICF constructs of impairment, activity limitation and participation restriction; and set 2 comprised 8 vignettes varying by severity of impairment and activity limitation, and level of patient motivation. Conjoint analyses were applied to determine the relative importance of the patient attributes and the part-worth values (utility) of each attribute level.

Results: Greater severity of impairment (pain), activity limitation and participation restriction and higher motivation were associated with a higher priority for surgery. In set 1, the most important attribute was pain (average importance = 52.2%), followed by activity limitation (25.5%) and participation restriction (22.3%), whereas in set 2, the most important attribute was patient motivation (43%), followed by pain (36%) and activity limitation (21%).

Conclusions: Although impairment had the greatest influence on prioritisation for THR in set 1, if the influence of activity limitation and participation restriction is combined into a single disability construct, then disability and impairment received similar weightings. This suggests a shift away from traditional impairment dominated models of disability towards more complex models containing behavioural constructs, such as activity limitation and participation restriction. Similarly, patient motivation had the greatest influence on prioritisation for THR in set 2, suggesting that complex psychological constructs also play a key role in clinical decision making around THR.

3.2 Introduction

Approximately 10% of people over the age of 60 in the Western world are said to have osteoarthritis of the hip (Dreinhofer et al., 2006) and this prevalence rate is estimated to double by the year 2020 (Badley & Crotty, 1995). Total hip replacement (THR) is an effective treatment for patients with moderate to severe arthritis of the hip, reducing pain and functional disability. Consequently, THR is one of the most frequently undertaken surgical procedures in orthopaedics (Birrell, Afzal, Nahit et al., 2002). Nevertheless, limited resources are characteristic of countries with publicly funded healthcare systems. Patients typically have to wait for THR and therefore, have to be assigned priority on a waiting list.

In Chapter 2, two clinical prioritisation tools for total joint replacement were examined to identify which factors of patient health and functioning, as defined by the International Classification of Functioning, Disability and Health (ICF) framework, were being assessed. The two tools were developed to manage waiting lists in New Zealand and Canada and other countries have also developed similar tools (Ebinesan, Sarai, Walley et al., 2006; Escobar, Gonzalez, Quintana et al., 2009). Nonetheless, the use of such tools is not standard practice in many countries, meaning that the clinician generally has to employ his or her own clinical decision making process to assign priority to patients for THR.

Wide variation in the indications for THR exists (Dreinhofer et al., 2006; Mancuso et al., 1996). However, pain, functional limitations and some radiographic changes have generally been identified as important clinical indicators that can influence clinicians' decision to offer THR (Birrell et al., 2002; Dreinhofer et al., 2006; Mancuso et al., 1996; National Insitute of Health Consensus Panel, 1995) and influence prioritisation for surgery (Dolin, de C Williams, Ashford et al., 2003; Glozier, Groom, & Prince, 2004). Although, these indicators have been generally accepted, the relative importance of each remains unclear. Further, there are no minimum criteria available for the clinician to use to determine whether a patient qualifies for THR or what priority for surgery they should be allocated (Dolin et al., 2003).

In contrast to the relatively strong support for the influential role of clinical indicators in decision making around THR, evidence demonstrating a role for patient psychological factors or characteristics is more limited. Mancuso et al (1996) found that a patient's lack of motivation swayed orthopaedic surgeons' decision against surgery, and a patient's desire to be independent and return to work swayed the surgeons' decision in favour of surgery. Thus, it could be argued that if these psychological factors have the potential to modify the surgeon's decision to offer surgery then they are likely to play a role in the allocation of priority for surgery. However, studies have shown that patients' expectation of improvement following surgery (Dolin et al., 2003), their psychological distress or illness perceptions (Glozier et al., 2004) were not associated with surgical priority.

To date, experimental studies which have investigated the indications for THR have asked participants to consider each possible indicator in isolation. For example, Dreinhofer et al (Dreinhofer et al., 2006) asked participants to select whether each indicator was of high, intermediate or low importance, whilst Mancuso et al (1996) asked participants to select from a list of possible responses the least severe level of an indicator that would be consistent with performing surgery. However, during a consultation with a patient, the clinician is unlikely to consider each indicator in isolation but rather consider all possible indicators simultaneously to determine whether the patient qualifies for surgery and what priority they should be allocated. Therefore, a more comprehensive approach is needed to investigate the importance of each factor simultaneously.

Conjoint analysis is a methodology that can be used to investigate prioritisation for THR by modifying possible indicators simultaneously. Originating from Conjoint Measurement Theory (Luce & Tukey, 1964), conjoint analysis is based on the premise that products or services are comprised of various attributes and that each attribute has several levels, each of which has a unique value or utility for the individual considering them. Conjoint analysis produces two different types of values for each attribute: the utility or ‘part-worth’ of each level of the attribute and the relative importance of the attribute (Green & Wind, 1975). Part-worth utility values can be summed to provide the overall utility value for the product or service, which reflects the individual’s preference for the product or service in question. The relative importance of the attribute is expressed in the form of a percentage, computed by dividing the utility range for each attribute by the sum of all utility ranges. Within the context of prioritisation of a patient for THR, the product or service is the priority assigned to the patient; the attributes are the clinical indicators or patient factors and; the levels of an attribute are the different severities or levels of the indicators or factors. For example, radiographic evidence of the extent of joint damage, the patient’s age, whether or not the patient lives alone, and whether or not the patient is a carer, are all potential attributes that contribute towards the prioritisation of a patient for THR surgery.

Conjoint analysis methodology is common to economics and marketing where it is used to examine consumer preferences; however, it has recently begun to be applied in studies of healthcare. For example, it has been used to investigate patient preferences for treatment or services (Fraenkel, Bodardus, & Wittink, 2001; Kellett, West, & Finlay, 2006; Ryan & Farrar, 2000; Singh, Cuttler, Shin et al., 1998; Stanek, Oates, McGhan et al., 2000); older adults' preferences for walking programmes (Brown, Finkelstein, Brown et al., 2009); healthcare professionals' decisions about treatments (Gunnarsdottir & Kinnear, 2005; Raley, Followwill, Zimet et al., 2004) and; the allocation of scarce medical resources (Furnham, Hassomal, & McClelland, 2002). Further, conjoint analyses have been applied to study the prioritisation of patients with varicose veins, hernia and gallstones to surgical waiting lists (Oudhoff, Timmermans, Knol et al., 2007). Nevertheless, to date, the methodology has not been employed in the study of prioritisation for THR.

In sum, there is a need for research to examine the factors influencing prioritisation for THR simultaneously. Moreover, previous research has not applied a theoretical framework to investigate the indications for THR or their influence on prioritisation for THR. The ICF is the main international model of health outcomes, therefore it is an appropriate theoretical framework to apply to clinical decision making. Chapter 2 applied the ICF framework to investigate what constructs were being measured by clinical prioritisation tools for total joint replacement. In this chapter, the ICF will be applied to investigate the clinical decision making of orthopaedic surgeons in the absence of such prioritisation tools and to identify the constructs that influence prioritisation for THR. The ICF contains constructs which are traditionally of interest to orthopaedic consultants such as impairment and activity limitation. It also contains personal factors, such as patient motivation that previous literature has identified as influencing decisions around the provision of THR surgery. Further, personal factors can include psychological constructs such as intention (motivation) taken from the Theory of Planned Behaviour (TPB), which feature in the integrated ICF/TPB model

(Johnston et al., 2002). Thus, this study will investigate constructs found in the ICF and the integrated ICF/TPB model.

3.2.1 Research Questions

1. What relative importance is afforded to the ICF constructs of impairment, activity limitation and participation restriction in prioritisation of patients for THR and what are the part-worth values for each level of a construct?

2. What relative importance is afforded to the ICF constructs of impairment and activity limitation, and to the construct of patient motivation in prioritisation of patients for THR and what are the part-worth values for each level of a construct?

3.3 Methodology

3.3.1 Design

A conjoint study using a full factorial design was employed. Orthopaedic surgeons ranked two sets of patient vignettes in order of priority for THR. The vignettes differed in terms of three patient attributes expressed at two different levels. The full-factorial design ensured all possible combinations of each attribute by each level were expressed.

3.3.2 Participants

A convenience sample of 55 orthopaedic surgeons attending the 2007 European Federation of National Associations of Orthopaedics and Traumatology congress were invited to take part in the study. Exclusion criteria were not being a fluent English speaker and reporting a profession other than orthopaedic surgeon (e.g., rheumatology nurse). Forty-two surgeons (39 male, 1 female and 2 who did not report gender) accepted the invitation and gave informed consent (76% participation rate). The mean age of participants was 43.9 years (range 27 to 61 years) with an average of 15 years since qualification (range 1 to 36 years). Ninety percent were of European nationality and 93% practiced in Europe. The average number of THR performed by the participant or their team in 2006 was 232 operations.

3.3.3 Measures and Materials

A full factorial design was used to generate two sets of patient vignettes. Three attributes, each at two levels, were systematically varied, thereby producing eight vignettes per set. Set 1 addressed research question 1 and set 2 addressed research question 2.

Set 1 consisted of the attributes impairment, activity limitation and participation restriction. The ICF classifies pain as impairment (Cieza et al., 2004; Dreinhofer et al., 2004) and DCV studies of pain measures consistently show that items assessing pain measure the impairment construct (Dixon et al., 2007; Pollard et al., 2006); therefore, in this study the ICF construct of impairment (I) was defined as *pain.* Activity limitation and participation restriction were defined using the ICF theoretical definitions: Activity limitation (A) *difficulties an individual may have in executing activities;* Participation restriction (P): *problems an individual may experience in involvement in life situations.* These definitions were provided below the vignettes for easy referral. Pain, A and P were expressed at two levels: ‘moderate’ and ‘severe’.

Set 2 consisted of the attributes of impairment (I), activity limitation (A) (defined as in set 1) and patient motivation. Patient motivation was not defined further. Pain and A were expressed at two levels: ‘moderate’ and ‘severe’. Patient motivation was expressed at two levels: ‘not motivated’ and ‘highly motivated’. The vignettes in each set were presented in a fixed but random order (see Table 3.1).

Table 3.1. The type and level of attributes expressed in each patient vignette for each set

	Set 1			**Set 2**		
Patient vignette	Pain	Activity limitation	Participation restriction	Pain	Activity limitation	Motivation
1	severe	severe	severe	severe	severe	not
2	moderate	moderate	severe	moderate	moderate	highly
3	severe	severe	moderate	severe	moderate	not
4	moderate	severe	severe	moderate	severe	not
5	severe	moderate	severe	severe	severe	highly
6	moderate	moderate	moderate	moderate	severe	highly
7	severe	moderate	moderate	severe	moderate	highly
8	moderate	severe	moderate	moderate	moderate	not

3.3.4 Procedure

Participants were provided with a questionnaire displaying the two sets of patient vignettes (see Appendix 3 for an illustration of the vignettes). Participants were informed that all patients were 65 years old with radiographic abnormalities of the left hip and that all had elected for THR. Participants were asked to rank each set of patient vignettes in order of priority for THR. They were directed to assign a ranking of 1 to the patient vignette they considered to be the highest priority, 2 to the vignette they considered to merit the second highest priority, and so on, finally assigning a ranking of 8 to the vignette they considered to have the lowest priority. This procedure was carried out for set 1 and then for set 2. Finally, participants completed a series of demographic questions and questions about their current clinical practice.

3.3.5 Analyses

Conjoint analyses investigated orthopaedic surgeons' prioritisation of patients for THR. For each attribute two indices were calculated: the part-worth utility of each attribute and the relative importance of the attribute within each set. Note, conjoint analyses produce results that are only valid within a particular set of vignettes; thus, the analyses do not support comparison of indices between sets.

Calculations were done for each participant and then averaged over the sample. A linear model was specified with the expectation that higher levels of an attribute would correspond with higher assigned priority. This *a priori* specification does not affect utility estimates but simply allows participants showing a variation from this expected relationship to be identified. Participants who failed to rank any of the vignettes were identified and excluded from analysis (N = 2 for set 1; N = 3 for set 2), whilst those who assigned equal priority to three or more vignettes were identified and these particular cases were ignored. Consequently, the final sample comprised 40 participants for set 1 and 39 participants for set 2. Econometric guidelines about the sample size needed when using conjoint analysis indicate that a sample size of 30-100 participants is sufficient depending on the experimental design (Pearmain, Swanson, Kroes et al, 1991). For investigational research of this design using conjoint analyses, the sample size was typical (Orme, 2006).

Kendall's concordance coefficient *W* statistic (Kendall, 1948) was computed to assess the concordance in rankings between participants. Unlike other reliability statistics, such as the intraclass correlation employed in Chapter 2, no recommendations or benchmark values exist to make inferences about the strength of Kendall's *W* value. The

W value can only be judged with reference to the guidance that a score of 0 indicates total disagreement and a score of 1 indicates total agreement.

3.3.6 Ethics

Ethical approval was obtained from the Department of Psychology, University of Stirling.

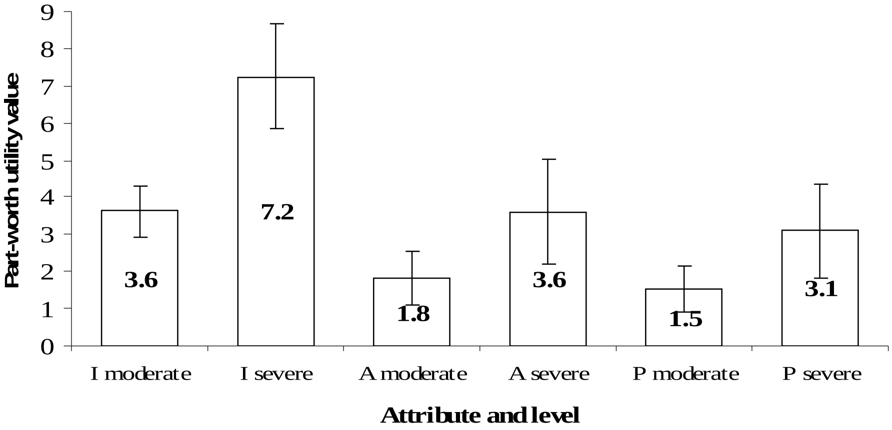
3.4 Results

3.4.1 Set 1:Impairment (I), Activity Limitation (A) & Participation Restriction (P)

Figure 3.1 displays the part-worth values for each level of the attributes expressed in set 1: I, A and P. Greater severity of each attribute is shown to be associated with higher part-worth scores, which indicates higher priority for THR. This concurs with the correlation between the actual rank order and the predicted rank order, as specified by the linear conjoint model (r=.99, p<.001). A positive relationship between the severity of an attribute and the part-worth score was identified in the rankings made by all participants. The concordance in rankings between participants was high (W=0.91); a score of 1 reflects complete agreement (Kendall, 1948).

The highest part-worth value was seen for ‘severe I’ (7.2), at least double the size of the values for ‘moderate I’ (3.6), ‘severe A’ (3.6) and ‘severe P’ (3.1). The part-worth value for ‘moderate I’ was equal to and greater than the values identified for ‘severe A’ and ‘severe P’ respectively.

Figure 3.1. Mean part-worth utility values for moderate and severe levels of impairment (I), activity limitation (A) and participation restriction (P).



Note. Bars represent standard deviations

The relative average importance values for each attribute in set 1 are presented in Table 3.2. Impairment was the most important attribute in the prioritisation of patients for THR in this set, assuming over half of the relative average importance.

Table 3.2. Relative average importance values (%) for the attributes expressed in each set.

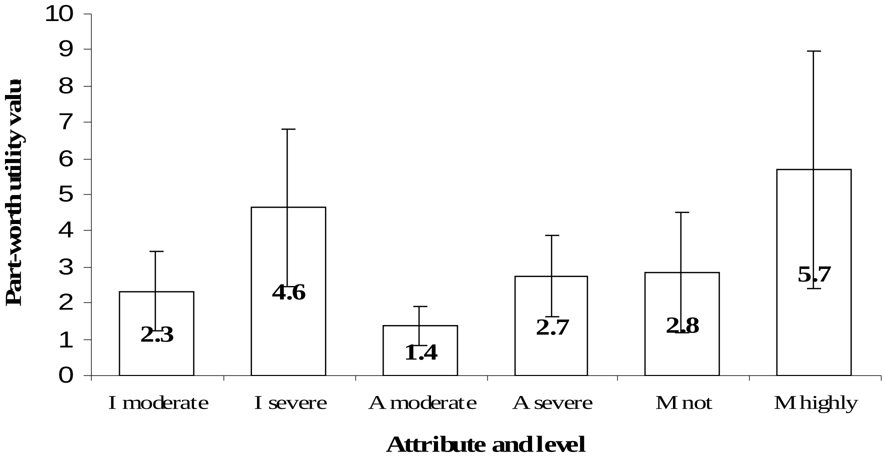
	Set 1			**Set 2**		
	I	A	P	I	A	M
Relative Average Importance %	52.2	25.6	22.2	35.6	21.4	43.0

3.4.2 Set 2 – Impairment (I), Activity Limitation (A) & Patient Motivation

Figure 3.2 displays the average part-worth values for each level of the attributes expressed in set 2: I, A and motivation. Higher levels of each attribute corresponded to higher assigned priority for THR, shown by higher part-worth scores. Within the conjoint model, the actual and predicted rank order were highly correlated (*r*=.99, *p*<.001). A positive linear relationship between the severity of an attribute and the part-worth score was observed in all but one case; data from one participant revealed a negative relationship between the level of motivation and part-worth score, indicating that a patient with no motivation is assigned higher priority than a highly motivated patient. Kendall's *W* statistic was 0.75 indicating the degree of concordance in rankings between participants was lower than observed for set 1.

The highest part-worth value was seen for the patient attribute of 'highly motivated' (5.7), followed by 'severe I' (4.6). The attributes of 'not motivated' and 'severe A' exhibited part-worth values approximately half that of 'highly motivated'. Similarly, 'moderate I' produced a part-worth value half that of 'severe I'. The lowest part-worth value was seen for 'moderate A'.

Fig. 3.2. Mean part-worth utility values for moderate and severe levels of impairment (I) and activity limitation (A), and no and high levels of motivation (M)



Note. Bars represent standard deviations

The relative average importance values for each attribute in set 2 are also presented in Table 3.2. Motivation was the most important attribute in the prioritisation of patients for THR in this set, assuming 43% of the relative average importance. This was double the relative importance assigned to A (21.4%) and greater than that assigned to I (35.6%).

3.5 Discussion

Conjoint analyses revealed that the ICF can be used as a theoretical framework to understand clinical decision making in relation to prioritisation for THR. The large majority of participants were willing and able to rank patient vignettes based on attributes derived from the ICF and the integrated ICF/TPB model, which explicitly defines psychological constructs.

Hypothetical patients described in vignette set 1 varied in three attributes based on the ICF constructs of I (operationalised as pain), A and P. Each ICF construct was considered in the prioritisation of patients for THR. However, the ICF constructs were not afforded equal importance: pain dominated the model. Surgeons assigned pain twice the relative importance of either A or P in their prioritisation decisions. The importance of pain in the prioritisation decision was emphasised further by the part-worth data. Pain, at both moderate and severe levels, was valued as a more important characteristic than either A or P. Indeed, a moderate level of pain was associated with a similar degree of priority as a severe level of A or of P. Pain, like many other symptoms of a health condition such as visual disturbances or shivering, is experiential in nature. In this study and in accordance with the general agreement in the ICF literature, including literature specific to osteoarthritis (Dreinhofer et al., 2004; Pollard et al., 2006), pain is also considered to be measure of impairment. Therefore, the current findings illustrate the pivotal role of a patient’s impairment in the clinical decision of prioritisation for THR.

That said, there is some discussion within the literature as to whether the concepts of A and P should be distinct or combined (Jette et al., 2007). Whilst the ICF presents A and P as distinct theoretical concepts, each with its own definition (World Health Organization, 2001), the work on the ICF core measurement sets has yet to produce discriminant measures of A and P (Cieza et al., 2004; Dreinhofer et al., 2004). For example, the core measurement set for osteoarthritis details the categories to measure I, but combines the measurement of A and P. In addition, investigation into whether osteoarthritis health outcome instruments can operationalise the ICF found that 11 of 13 instruments contained items measuring a combined AP health outcome domain (Pollard et al., 2006). Thus, although it is possible to measure A and P with discriminant validity (Dixon et al., 2007; Dixon et al., 2008a; Pollard, Dixon, Dieppe et al., 2009), the issue of whether A and P should be combined into a single disability construct remains a topic of active discussion.

Both A and P are disability behaviours; A is defined as limitations in the execution of activities, and P is defined as restrictions in involvement in life situations. Therefore, there would appear to be no conceptual problem with combining the constructs of A and P to create a single behavioural construct of disability, namely activity limitation & participation restriction (AP). Based on the summation of the identified relative importance of A and P in the current study, the combined construct of AP would have had a relative importance of 48%, a value not dissimilar to the 52% identified for I. However, conjoint analysis produces relative importance values (i.e., the importance of one attribute in relation to the other attributes being examined), therefore, the relative importance of I and a combined attribute of AP would need to be tested directly by examining the importance of I in direct comparison with a combined AP construct. The possibility that orthopaedic surgeons place almost equal importance on the behavioural attribute of AP as they do on the medical attribute of I, in the assignment of priority to patients for THR, suggests a shift away from traditional medical models of health and illness, which emphasise biological functioning towards more complex models able to accommodate factors that influence behaviour, including disability behaviour.

The second research question investigated whether a patient's motivation influences the priority assigned to them for THR, and what relative importance is afforded to motivation when judged alongside attributes of I (measured as pain) and A. This question was addressed by the hypothetical patients in vignette set 2. Patient motivation was identified as having greater relative importance than either pain (I) or A. In addition, high patient motivation had the highest part-worth utility score when compared to all attributes at all levels i.e., high motivation was of greater importance than either severe pain or severe A. This finding is consistent with the finding of Mancuso et al (1996), who reported that a lack of patient motivation swayed orthopaedic surgeons' decision against THR. These data suggest that surgeons' perceptions of patient motivation influence important aspects of their decision making around THR. Motivated patients may be more likely to be offered THR in the first place and may then be assigned higher priority on the waiting list.

The current study utilised hypothetical patients to simulate the decision making process of orthopaedic surgeons when assigning priority to patients for THR. It did not ask surgeons to provide the reason for their decisions; therefore we can only hypothesise why being highly motivated was associated with greater priority. Unlike the ICF constructs, no theoretical definition for the attribute of patient motivation was given to participants. Consequently, participants may have interpreted the concept of patient motivation in many different ways. For example, high motivation may have been interpreted as an indication of patients who are more likely to attend post-surgery rehabilitation or adhere to medication regimes, or perhaps an indication of those less likely to fail to attend for surgery.

However, notwithstanding this concern, reliability analyses of set 2 indicated that the participants ranked the vignettes similarly which implies that, regardless of how patient motivation was interpreted, the attribute was consistently influential. For example, internal consistency and reliability of the conjoint model was high (i.e., the actual and predicted rank order was highly correlated) and; the value reflecting concordance in rankings between participants (*W*=0.75) was closer to a value of 1 (total agreement) than a value of 0 (total disagreement). It is noted that concordance for set 2 (*W*=0.75) was lower than identified for set 1 (*W*=0.91). Set 2 differed from set 1 by the inclusion of the construct of patient motivation rather than the construct of P, therefore the lower concordance in rankings between participants in set 2 may reflect different interpretations of the patient motivation attribute and different judgements regarding its relative importance in prioritisation for THR. Furthermore, in analyses of set 2, one participant was identified as ranking patient motivation in the opposite direction to other participants, i.e., a patient with no motivation was assigned higher priority than a highly motivated patient. Removal of this case from analyses resulted in an increase in the concordance coefficient from 0.75 to 0.79, suggesting that this participant’s data did affect concordance between participants. However, even after removal of this case, the concordance value was still lower than identified for set 1. Further investigation of how patient motivation influences the prioritisation decision and how this influence varies between participants is needed.

There is a plethora of evidence in support of an association between clinical indicators, such as pain and functional limitations, and a clinician’s decision to offer THR and prioritisation for surgery (Dolin et al., 2003; Dreinhofer et al., 2006; Glozier et al., 2004; Mancuso et al., 1996). However, the evidence for the influential role of patient psychological factors is mixed; patients’ expectation of improvement (Dolin et al.,

2003) and psychological distress and illness perceptions (Glozier et al., 2004) were not found to be associated with surgical priority, whilst a patient’s lack of motivation swayed orthopaedic surgeons’ decision against surgery, and a patient’s desire to be independent and return to work swayed the surgeons’ decision in favour of surgery (Mancuso et al., 1996).

The current study similarly found that the psychological factor of patient motivation was an important factor in surgeons’ clinical decisions regarding THR; however unlike Mancuso et al’s (1996) study, which examined surgeons’ decision to offer surgery, the current study examined surgeons’ prioritisation of patients for THR. It is possible that patient motivation but not psychological factors such as expectations about surgery, distress and illness perceptions is afforded importance by surgeons because motivation is interpreted in terms of its behavioural consequences. For example, as discussed above, high motivation may be interpreted as an indication of patients who are more likely to adhere to post-surgery rehabilitation or medication regimes. Qualitative exploration is needed to elucidate the precise meaning or meanings of patient motivation for orthopaedic surgeons in relation to the decision to offer surgery and subsequent prioritisation for THR.

The finding that a psychological construct, such as patient motivation, was not only important in the prioritisation decision but actually more important than the clinical factors of pain and A, provides strong support for the integration of psychological models into the ICF as proposed by Johnston & Pollard (2001). The current findings offer preliminary evidence that orthopaedic surgeons may use components of the integrated ICF/TPB model in their decision making about priority for THR.

Nevertheless, the explicit rationale behind their decision to assign higher priority to patients who are highly motivated is yet to be discerned.

Conjoint analyses has provided useful insight into the relative importance assigned to various patient attributes and the trade-offs made by surgeons when required to simultaneously judge three patient attributes. However, a weakness of the technique lies in the fact that data are derived on the basis of individuals' judgements about hypothetical situations and therefore criterion validity of the findings needs to be established. This study asked orthopaedic surgeons to assign priority to paper-based patient vignettes and therefore, it is acknowledged that the judgements reported in this study may differ from those made in an actual clinical situation. However, a review of conjoint analysis studies concluded that studies which are both rigorously designed and analysed can predict real behaviour (Louviere, 1988 cited in Ryan, 1996), lending some support to the validity of the findings.

The ICF framework used in this study provides additional potential to evaluate the validity of the findings in this study. Application of the ICF core measurement set for osteoarthritis (Dreinhofer et al., 2004) would enable patients on waiting lists for THR to be assessed in terms of their relative levels of I and AP. It would be possible to identify whether a patient's I (pain) was given a similar importance to their AP in determining the priority with which they are assigned as was found in set 1. Equally, utilisation of the integrated ICF/TPB model would permit exploration of the relative importance of ICF constructs and psychological characteristics, such as patient motivation, shown to be important in the findings of set 2.

A general limitation of conjoint analysis studies is that the number of attributes and levels studied are restricted by the inability of participants to judge an infinite number of attributes and therefore only a selection of attributes and levels can be considered at any one time. In reality orthopaedic surgeons may have extensive information about each patient from which they base their decision. In the current study only three attributes were examined in each set of vignettes and this simplification is likely to have affected the results. For example, important factors such as whether the patient lives alone were not included. The number of factors that may potentially influence selection for and then prioritisation for THR is large; increasing the number of attributes and the number of attribute levels within a conjoint study rapidly inflates the number of vignettes participants are required to rank. The full factorial design used in the current study would need to be replaced by a fractional factorial design. A fractional factorial design reduces the number of vignettes required; only those vignettes necessary to test orthogonal main effects of the attributes of interest are ranked by participants. In this way, a fractional factorial design would enable several attributes at several different levels to be assessed simultaneously without significantly increasing the burden on participants in terms of the number of vignettes they are required to rank.

This study examined the clinical decision making for the prioritisation of patients for THR by orthopaedic surgeons. In clinical practice, before a patient is seen by an orthopaedic surgeon, they are typically seen by a referring physician who acts as a gatekeeper to access to surgery. Dreinhofer et al (2006) found that referring physicians and surgeons differ in their approach to indications for recommending THR; therefore, it seems plausible that if the current study were replicated with a sample of referring physicians, different results may be found. As noted by Dreinhofer et al (2006), in order to reduce variations and possible inequities in the provision of care, objective indication criteria are needed. Further investigation into the criteria used by clinicians at the referral and prioritisation stage is needed; the use of patient vignettes within a conjoint analysis methodology is a valuable tool to approximate current practice.

Finally, due to the convenience sample and exploratory nature of this study a relatively small sample size was used. A larger sample size would allow prioritisation patterns to be evaluated among sub-populations of clinicians. For example, aggregate mean part-worth utilities and importance could be examined for surgeons practicing in different countries with different healthcare systems to see whether predictors of prioritisation exist within and between various sub-populations. This may help to explain the international differences in THR provision and even inform future healthcare strategies.

3.5.1 Summary and Implications

Orthopaedic surgeons show concordance of judgements regarding the relative importance of attributes in the clinical decision of prioritisation of patients for THR. Evidence from this study suggests that they draw on information from both medical and psychological models of health and illness, placing similar importance on impairment (pain) as on a combined behavioural construct of AP. Furthermore, when patient motivation is considered in the judgement, it is afforded greater importance than either I or A. These findings lend support to the integrated ICF/TPB model which integrates psychological variables into the ICF framework, in order to further understand this important decision making process.

Chapter 4 - Predicting recovery after total hip replacement – the role of pre-operative patient expectations

4.1 Abstract

Background: Pain and functional limitations can be significantly reduced by total hip replacement surgery (THR). However, variability in the success of surgery exists with some patients reporting residual pain and disability after surgery. Patient expectations have been identified as a factor that may account for individual differences in recovery after surgery. This study employed the International Classification of Functioning, Disability and Health (ICF) to classify pre-operative patient expectations and explore the relationship between patient expectations and recovery after THR.

Method: A European cohort of 1108 patients reported two types of expectations of THR as part of their pre-operative assessment. Patients reported what they anticipated surgery would enable them to do, that they needed to be able to do, but that they could not currently do (‘need’ expectation). In addition, patients reported what they would like to be able to do in a year’s time that they currently could not do (‘desire’ expectation). One year post-surgery, patients’ health and functioning were reassessed (WOMAC & EQ-5D) and patients reported what they could do now that they could not do prior to surgery. Free-text responses to the pre-operative expectation questions and the 1-year post-surgery current function question were classified to the ICF constructs of impairment, activity limitation and participation restriction by two researchers. Concordance between at least one of the ICF classified pre-operative expectations and ICF classified post-operative function question identified patients whose expectations had been met. Analyses explored the relationship between the ICF classified expectations and post-operative recovery.

Results: All patient expectations were classified to the ICF. Less than 5% of patient expectations were identified as impairment, 58% of ‘need’ expectations were identified as activity limitations and 45% of ‘desire’ expectations were identified as activity limitations & participation restrictions combined. After controlling for demographic and clinical variables, reporting more ‘need’ expectations was associated with better pain-related recovery (WOMAC pain subscale). However, no other significant relationships were identified between the number of patient expectations and any measure of recovery from surgery. Further, no significant differences in recovery were found on the basis of the content of patient expectations, e.g., patients expressing activity limitation expectations did not make a better recovery than patients expressing activity limitation & participation restriction expectations. Finally, better recovery from health-related quality of life (EQ-5D) was associated with having ‘met’ expectations.

Discussion: THR targets impairment, however few patient expectations were classified to the ICF definition of impairment. The majority of patient expectations were classified as activity limitation or a combination of activity limitation & participation restriction. Thus, patient expectations of surgery, as measured in this study, focus on recovering valued activities rather than reversal of bodily impairments. There was very limited evidence of a relationship between the number or content of pre-operative expectations and post-operative recovery. However, this null finding may be attributable to the nature of the pre-operative expectation questions that were asked. The need for additional studies in this area is discussed.

4.2 Introduction

In the Western world, the prevalence of hip osteoarthritis is high with 10% of people aged 60 years and older reporting symptoms (Dreinhofer et al., 2006). People with osteoarthritis frequently report pain, and severe limitations in their functional ability and their ability to perform social roles (Dekker et al., 1992). Total hip replacement surgery (THR) is an effective and frequently performed procedure for people with severe osteoarthritis of the hip, producing measurable reductions in pain and disability, and improvements in quality of life (Ethgen et al., 2004; Learmonth, Young, & Rorabeck, 2007; National Joint Registry, 2007; Orbell, Espley, Johnston et al., 1998; Rissanen, Aro, Slatis et al., 1995). Despite the general effectiveness of THR, variability in the success of surgery exists. For example, figures show that approximately 10% of patients do not experience considerable long-term improvements in pain and function following THR (Fortin et al., 1999). Further, a systematic review of 118 studies of THR concluded that between 16 and 57% of patients report residual pain following THR (Fitzpatrick, 1988, cited in Zhang et al., 2005). This identified variability in the success of THR raises the question of what factors may account for individual differences in outcome and recovery.

Jones et al (2007) reviewed the literature on patient outcomes after THR and total knee replacement (TKR) and identified the following potential determinants of outcome: perioperative surgical complications, prosthetic-related factors and nonsurgical medical factors including preoperative pain and function, obesity and comorbid conditions. Age did not independently affect pain and functional outcomes; however, it was noted that older age is associated with more comorbid conditions and therefore may be indirectly associated with poorer outcome. Fortin et al (1999) examined the effect of patient demographic and clinical variables on pain and function 6-months after THR or TKR and concluded that the orthopaedic referral centre, education and comorbidity were important predictors of outcome but that pre-operative pain and function was the single best predictor. Despite the identified associations between demographic and clinical factors and outcome after surgery, these factors explained less than 28% of the variance in outcome after THR, meaning that a large percentage of variance in outcome remained unexplained.

Affective and cognitive psychological factors have also been identified as potential determinants of outcome and recovery from various types of surgery including THR (Jones et al., 2007; Kopp et al., 2003). Slower recovery after surgery has been associated with high levels of neuroticism or trait anxiety (Mathews & Ridgeway, 1981) and pre-operative state anxiety (Kopp et al., 2003), whereas faster recovery has been associated with dispositional optimism (i.e., the general expectation that good, as opposed to bad, outcomes will occur) (Scheier, Matthews, Owens et al., 1989). It is thought that these psychological factors may affect recovery through their influence on and association with coping behaviours (Kopp et al., 2003; Scheier et al., 1989).

Specific beliefs, namely patient expectations, have received much attention for their role as predictors of surgical outcome and recovery. Uhlmann, Inui & Carter (1984) defined patient expectations as expectancies or perceptions that “given events are likely to occur during or as a result of medical care” (pp.681). The evidence for a relationship between patient expectations and surgical outcomes is mixed. Some studies indicate that positive expectations predict better recovery and outcome after surgery, after controlling for clinical and demographic factors. For example, a higher level of expectation regarding positive results of treatment predicted better improvement and performance 1-year after surgery to repair rotator cuff tears (Henn, Tashjian, Kang et al., 2007); more patients with favourable expectations of surgery for sciatica reported good post-surgery outcomes than patients with unfavourable expectations (Lutz, Butzlaff, Atlas et al., 1999) and; patients who did not expect to experience post-operative pain following lumbar surgery were less disappointed than those who expected to experience post-operative pain (de Groot, Boeke, & Passchier, 1999). Moreover, specific to the joint replacement literature, patients with more positive expectations were less depressed 9-months after surgery (Orbell et al., 1998) and the expectation of complete pain relief independently predicted 6-month post-surgery function and improvement in pain levels (Mahomed, Liang, Cook et al., 2002).

However, other studies have failed to find a relationship between patient expectations and recovery after surgery, after controlling for clinical and demographic factors. For example, positive expectations did not predict post-operative symptoms or general health in patients who had undergone prostate surgery (Flood, Lorence, Ding et al., 1993) nor did they predict outcome or satisfaction in patients 2-years after TKR (Mannion, Kampfen, Munzinger et al., 2009).

Mondloch, Cole & Frank (2001) conducted a systematic review of the evidence for a relationship between patient recovery expectations and health outcomes and concluded that 15 of the 16 articles considered by the review provide support for a relationship between positive expectations and better health outcomes. However, this review was not specific to surgical interventions, but included studies of any medical intervention. Thus, it cannot be inferred that the predictive role of patient expectations is maintained in relation to surgical interventions and recovery.

In summary, evidence for a relationship between patient expectations and recovery after surgery is inconclusive. Reasons for the inconclusive findings may reflect the different surgical procedures or patient populations under investigation in each study; or alternatively, the variations in study methodology, such as the conceptualisation, elicitation and measurement of patient expectations. No standardised method of measuring patient expectations is available (Uhlmann, Inui, & Carter, 1984) and consequently, patient expectations have been measured in several ways. For example, studies assessing pain-related expectations of orthopaedic surgeries have varied significantly in terms of the wording of questions and the measurement scale employed. For instance, some studies have assessed *future pain* using set response categories, such as ‘none at all’, ‘much less’, ‘slightly less’ or ‘not altered’ (Eisler, Svensson, Tengstrom et al., 2002), whilst others have employed 4-point Likert scales to assess *pain relief* (‘no pain’ to ‘very painful’) (Mahomed et al., 2002) or 5-point Likert scales to assess *relief from symptoms (pain, stiffness etc.)* (‘not at all likely’ to ‘extremely likely’) (Henn et al., 2007). In contrast, other studies have used an open-ended approach to elicit patient expectations of surgery (Mancuso, Salvati, Johanson et al., 1997). This approach does not constrict participants’ responses, however, it does lead to another important measurement decision regarding how to group or classify free-text responses.

Another limitation of the literature in this area is that only a few studies have applied a theoretical framework to guide their investigation. Bandura’s self-efficacy theory (1977) is one theoretical framework that has been applied to investigate and interpret findings regarding the role of patient expectations on general health outcomes and recovery. Self-efficacy has been cited as playing a pivotal role in determining an individual’s recovery expectations; previous mastery experiences, vicarious learning and verbal persuasion have all been highlighted as possible contributing factors

(Mondloch, Cole, & Frank, 2001). Leventhal's self-regulatory model of illness (Leventhal, Leventhal, & Contrada, 1998) has also been used to explore the relationship between patient beliefs, and health outcomes and recovery (Horne, 1997; Petrie, Weinman, Sharpe et al., 1996; Orbell et al., 1998). The self-regulatory model states that an individual's illness representation, comprising component beliefs about the *identity*, *causes*, *timeline*, *consequences* and *cure/controllability* of the illness, is used to guide coping behaviour. This framework examines illness beliefs rather than specific patient expectations of recovery; however, both illness beliefs and patient expectations have been shown to predict recovery, therefore, it is possible that they are related concepts. Indeed, it has recently been proposed that the *consequences* component of illness representations from the self-regulatory model is conceptually similar to the concept of outcome expectations from self-efficacy theory (Lau-Walker, 2006).

In sum, few studies have applied theoretical frameworks to investigate the relationship between patient expectations and health outcomes, and those which have, have principally used the theory to help interpret their findings in accordance with the view that patient expectations/beliefs can play a key role in determining an individual's recovery. No studies have used a theoretical framework to define and classify patient expectations, thus, the aforementioned problem relating to the measurement of patient expectations persists. Applying a theoretical framework to define and classify patient expectations and to further investigate the relationship between expectations and recovery after surgery is needed to facilitate the development of a cumulative evidence base within the patient expectation literature.

The International Classification of Functioning, Disability and Health (ICF) model of health outcomes (World Health Organization, 2001) is a suitable theoretical framework to apply to the development of a cumulative evidence base in this area. The model's three health outcomes of disease, namely impairment (I), activity limitation (A) and participation restriction (P), can be applied to classify patient expectations (Beth Pollard, personal communication). The relationship between ICF classified patient expectations and health outcome can then be explored.

The current study investigates patient expectations of THR. The ICF framework is applied to investigate the relationship between patient expectations and recovery after surgery. THR is an elective procedure typically conducted at the later, more severe stages, of osteoarthritis, therefore, patients will generally be very familiar with seeking healthcare advice (Ross, Sinacore, Stiers et al., 1990). Thus, during the decision making process regarding whether to opt for surgery or not, THR patients are likely to have developed expectations about the effectiveness of surgery and typical recovery trajectory. This study utilises data from a large European cohort of THR patients ('EUROHIP' consortium).

4.2.1 'EUROHIP' Consortium

The 'EUROHIP' consortium includes 20 orthopaedic centres in 12 different European countries. In 2002, the group agreed to investigate the indications for THR in the participating centres (Dreinhofer et al., 2006; Sturmer, Dreinhofer, Grober-Gratz et al., 2005) and develop a large cohort of patients undergoing primary THR for osteoarthritis. This cohort originally included a total of 1520 patients: 193 cases have since been removed from the cohort due to protocol violations. The final cohort comprises 1327 patients and, to date, has been used to examine the amount of variation in disease status (impairment and function) at the time of THR and to explore the possible clinical and demographic determinants of this variation (Dieppe, Judge, Williams et al., 2009). The relationship between pre-operative variation in disease status, clinical and demographic determinants, and 1-year post-surgery health outcomes is currently being investigated. Patient expectations of the operation were also collected pre-operatively and will be the focus of the analyses reported here.

4.2.2 Research Questions

1. Are patient expectations of THR consistent with the definitions of I, A and P constructs of the ICF?
2. Is there a relationship between the number of expectations expressed by patients and their recovery from THR?
3. Is there a relationship between the content of expectations expressed by patients and their recovery from THR? (i.e., are I expectations associated with recovery in a different manner to A expectations or P expectations?)
4. Is there an association between a patient's recovery and having their expectations met?

4.3 Methodology

4.3.1 Design

This is a prospective study of the 'EUROHIP' cohort of patients undergoing primary THR, which ran from 2002 to 2006. Demographic and clinical variables, disability, health status, and patient expectations of THR were measured prior to surgery. One-year post-surgery, disability and health status were reassessed and patient-reported functional outcomes were measured. The study measures were predetermined by the 'EUROHIP' consortium.

4.3.2 Participants

The final 'EUROHIP' cohort comprises a total of 1327 patients from the 20 participating orthopaedic centres. A minimum of 50 consecutive, consenting patients entered the study from each orthopaedic centre. Inclusion criteria for patients included a diagnosis of osteoarthritis of the hip, primary THR (i.e., no revision operations) and a willingness and ability to take part in the study. Exclusion criteria included reasons for surgery other than osteoarthritis, severe mental illness or dementia. Demographic and pre-operative clinical measures of this sample have previously been reported (Dieppe et al., 2009).

Only 1108 of 1327 patients in the cohort (83%) responded to one or other of the questions that elicited patient pre-operative expectations of THR; therefore the data reported in the current study is taken from this subsample of 1108 patients. Multivariate analysis of variance revealed that responders did not significantly differ to non-responders on any of the demographic or clinical variables (F (10, 728) = 1.13, p = .34). Six hundred and twenty-nine patients were female (57%). The mean age was 65.1 years (SD 11.0, range 25-91) and mean body mass index (BMI) was 27.6 (SD 26.9, range 18.6-47.9). Fifty-one percent of the sample had no qualifications after leaving school, 32% had a college diploma and 18% had university qualifications (see Table 4.2 for all demographic and clinical characteristics of the sample).

4.3.3 Measures

4.3.3.1 Disability

Disability was assessed using the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC: Bellamy, Watson Buchanan, Goldsmith et al., 1988). The WOMAC is a validated self-administered questionnaire used to assess symptom severity and disease-specific health-related quality of life in patients with osteoarthritis of the knee or hip. The index consists of 24 items grouped into three subscales: pain (5 items), stiffness (2 items) and physical function (17 items). Items are measured on a 5-point Likert scale (0 = none, 1 = slight, 2 = moderate, 3 = severe, and 4 = extreme); higher scores indicating greater symptom severity. Missing data were treated in accordance with previous data analyses of the 'EUROHIP' cohort (Dieppe et al., 2009): when ≥2 pain items, both stiffness items, or ≥4 function items were missing, the subscale was not calculated; when 1 pain item, 1 stiffness item, or 1-3 function items were missing, the average value for the subscale was used to replace the missing item(s). Scores from the items of each subscale were summed to create total scores for each subscale. These subscale scores were then summed to create a Total WOMAC score. Each of the three subscale scores and the Total WOMAC score were converted into normalised scores (0 = no symptoms and 100 = extreme symptoms) by multiplying them by 100 and dividing by the possible maximum score for the scale.

4.3.3.2 Health status and health-related quality of life

The EQ-5D (The EuroQol Group, 1990) descriptive system was used as a standardised, non-disease specific measure of health status and health-related quality of life. Concurrent use of a generic instrument, such as the EQ-5D, and a disease-specific instrument, such as the WOMAC used in this study, has been recommended in outcome studies of the elderly where comorbidity is likely (Hawker, Melfi, Paul et al., 1995; Quintana, Escobar, Bilbao et al., 2005).

The EQ-5D consists of 5 dimensions: mobility, self-care, usual activities, pain/discomfort and anxiety/depression. Each dimension has 3 levels (1 = no problems, 2 = some problems, and 3 = severe problems). These values do not have arithmetic properties but rather are combined to create a 5-digit number that reflects the respondent's health state based on the 5 assessed dimensions. A single index value for each of the 243 (3^5) possible health states was calculated using the valuation model developed by Dolan (1997). Dolan elicited direct valuations of a subset of the EQ-5D health states using the time trade-off method from a sample in the UK. The direct valuations were then interpolated using regression analyses to predict valuations for all possible EQ-5D states. The single index score for each health state ranges from -0.594 to +1, with 'dead' anchored at 0 and 'full health' at +1, and health states worse than dead at <1.

4.3.3.3 Recovery

Recovery variables were computed for the total WOMAC scale, each of the WOMAC subscales of pain, stiffness and physical function, and the EQ-5D index. Recovery was defined as the deviation of 1-year post-operative scores from the statistically expected scores derived from pre-operative scores. The recovery variables are the standardised residuals calculated from the regression of 1-year post-operative scores on the pre-operative scores. This method is commonly used to create an index of recovery within the disability and rehabilitation literature (Johnston, Morrison, Macwalter et al., 1999; Johnston et al., 2007; Johnston et al., 2004; Molloy, Sniehotta, & Johnston, 2009). Where necessary, recovery scores were transformed so that scores greater than 0 indicate better than average recovery, when based on the performance of the total group; and recovery scores less than 0 indicate worse than average recovery.

The regression equations used to create recovery variables showed that pre-operative Total WOMAC accounted for 13% of the variance in 1-year post-operative Total WOMAC (R=0.36, p <.001); pre-operative WOMAC pain accounted for 9% of the variance in 1-year post-operative WOMAC pain (R=0.30, p <.001); pre-operative WOMAC stiffness accounted for 3% of the variance in 1-year post-operative WOMAC stiffness (R=0.18, p <.001); pre-operative WOMAC physical function accounted for 15% of the variance in 1-year post-operative WOMAC physical function (R=0.38, p <.001); and pre-operative EQ-5D accounted for 9% of the variance in 1-year post-operative EQ-5D (R=0.29, p <.001). The range of residualised recovery scores for each measure were: -2.19 – 3.36 Total WOMAC; -1.70 – 3.96 WOMAC pain; -1.51 – 2.94 WOMAC stiffness; -2.23 – 3.23 WOMAC physical function; and -4.48 – 1.66 EQ-5D.

4.3.3.4 Patient expectations

Patient expectations of THR were elicited by two free-text response questions, one addressing ‘needs’ and the other addressing ‘desires’.

“What things do you think you might be able to do in a year’s time that you NEED to be able to do, but CANNOT do now, if the operation is a total success?” (‘Need’)

“What things do you think you might LIKE to be able to do in a year’s time that you CANNOT do now, if the operation is a total success?” (‘Desire’)

Patients’ responses to the expectation questions were quantified and classified using the following protocol. Firstly, two researchers independently counted the number of expectations given by each patient for each question. To identify distinct expectations the following rules were adopted: punctuation or ‘and’ operations between different concepts or actions indicated distinct expectations (e.g., the response *“painfree, more independent”* and the response “*to walk without problems and pain”* both counted as two expectations); however, when the operation ‘and’ joined two aspects of the same action then only one expectation was counted (e.g., “*to put on shoes and socks*”); when the response read “*nothing*”, the number of expectations was counted as zero; and when no response was given the data were treated as missing and the participant was classified as a ‘non-responder’ and not included in further analyses.

Secondly, the two researchers independently classified each of the first three expectations given in response to each question to the definitions of the ICF constructs of I, A and P (see Table 4.1). Both researchers were academics in health psychology and very familiar with the ICF framework. Each expectation was classified as either I, A, P, impairment & activity limitation (IA), impairment & participation restriction (IP), activity limitation & participation restriction (AP), or impairment & activity limitation & participation restriction (IAP). Mixed coding (IA, IP, AP, IAP) allow for concepts that clearly measured more than one ICF construct (e.g., “*to walk without pain*” was classified as IA). Responses that read *“nothing”* and responses that were not an expectation (e.g., “*I don’t know*” and “*I’d like to have the other hip done too*”) were treated as missing data and the patient was not included in further analyses relating to the content of that response. Similarly, patients providing only one expectation were treated as missing in analyses of the content of the second and third expectations, and patients providing two expectations were treated as missing in analyses of the content of the third expectation.

Table 4.1 Definitions of the WHO-ICF constructs

ICF Construct	Definition
Impairment (I)	Problems in body function or structures such as a significant deviation or loss
Activity limitations (A)	Difficulties an individual may have in executing activities
Participation restrictions (P)	Problems an individual may experience in involvement in life situations

4.3.3.5 Met or unmet patient expectations

As part of the 1-year post-operative questionnaire, patients were asked about their current functional outcomes with the following free-text response question:

"What things CAN you do now that you could not do a year ago, as a result of your hip operation?"

Using the same protocol used to classify the content of the responses to the pre-operative expectation questions, the two researchers independently classified each of the first three responses to the current function question to one of the ICF constructs or a combination thereof.

A dummy variable was created to indicate whether a patient's pre-operative expectations had been met or unmet. Having met expectations was determined by the ICF classification of at least one of the pre-operative expectations matching the ICF classification of any one of the responses to the 1-year post-operative current function question. For example, if a patient expressed at least one A expectation in response to either the 'need' or desire' pre-operative question, and then expressed at least one A response to the post-operative function question he or she was coded as having their expectations met. Having unmet expectations was determined by none of the ICF classifications of pre-operative expectations matching any of the ICF classifications of the responses to the 1-year post-operative current function question. This criterion included when a patient expressed a pre-operative *mixed* expectation that, in part, matched the ICF classification of a post-operative *pure* response. For instance, a patient expressing a *mixed* AP expectation and a *pure* A post-operative response was coded as having unmet expectations). The code of 8 was used to denote when a patient's response to the current function question read *"nothing"* and these patients were included in the 'unmet' expectation group. Patients defined as having 'met' expectations were coded 1 and patients defined as having 'unmet' expectations were coded 0.

4.3.3.6 Demographic and clinical variables

Demographic and clinical data were collected pre-operatively. Demographic data were self-reported in the pre-operative questionnaire including gender, date of birth (from which age was calculated), employment status (employed, retired, retired early or other), and education since leaving school (none, college diploma, university degree, postgraduate qualification). Clinical data were collected by the surgical teams including the patient's height and weight (from which BMI was calculated), side of surgery, duration of arthritis, date wait-listed, date of surgery and prosthesis type. The American Society of Anesthesiology comorbidity score (ASA: American Society of Anesthesiology, 1963) was also measured, which is a standard measure of fitness for surgery scored from 1 (normal, healthy) to 4 (life-threatening systemic disease). Finally, Kellgren and Lawrence (1957) radiographic scores were recorded as an indication of articular changes, ranging from 0 (no signs) to 4 (no joint space indicating severe changes).

4.3.4 Procedure

Questionnaires were piloted in Bristol, UK, modified accordingly and then translated for distribution to each European centre. Each centre recruited patients undergoing THR, who were willing and able to complete the self-administered questionnaires. Pre-operative questionnaires were completed in hospital prior to surgery. The 1-year post-surgery questionnaires were sent in the post to the patients. Responses to free-text response questions were translated from the patient's native language into English by bilingual representatives from each orthopaedic centre. Patient and centre anonymity was upheld with unique identifiers and the database was maintained and cleaned in Bristol. Centres were given 18 months to collect data.

4.3.5 Analyses

Firstly, attrition analyses were conducted to compare the final sample of those patients remaining in the study cohort 1-year after surgery with those in the study pre-operatively. Recovery variables were created using regression analyses. When either pre- or post-operative WOMAC or EQ-5D scores were missing, the recovery variable for that measure was not calculated and the case was excluded from all analyses relating to that recovery variable. Descriptive statistics of all study variables of the final cohort were calculated. Correlations between recovery variables, number of expectations and demographic and clinical variables were calculated. Demographic and clinical variables that could potentially confound the relationship between patient expectations and recovery were controlled for in analyses. This included age and gender as commonly controlled for demographic variables, educational qualification as an indication of socioeconomic status, BMI and ASA status.

All data were examined for outliers, normality, linearity and homoscedasticity. Post-operative WOMAC scores suffered from floor effects with many patients reporting a lack of disability after surgery producing positively skewed distribution curves for the WOMAC recovery variables. Similarly the post-operative EQ-5D scores suffered from ceiling effects with many patients reporting the best possible health-related quality of life after surgery, resulting in a negatively skewed distribution curve for the EQ-5D recovery variable. Transformation of these recovery variables did not normalise the distribution. Examination of the data for outliers suggested a small number of cases with recovery scores different to the majority; however, calculation of Mahalanobis distances indicated that these cases were not sufficient to warrant their removal from analyses when compared to critical values for each of the *a priori* analyses (Tabachnick & Fidell, 2001). Heteroscedasticity was evident in the recovery variables and is likely to be a result of the skewed distributions of post-operative scores used to create the recovery variable. However, heteroscedasticity is said to weaken rather than invalidate analyses (Tabachnick & Fidell, 2001).

Research question 1: Are patient expectations of THR consistent with the definitions of I, A and P constructs of the ICF?

This question was addressed by the classification of patient expectations to the definition of the ICF constructs. The Kappa statistic was employed to assess the level of agreement between the two judges.

Research question 2: Is there a relationship between the number of expectations expressed by patients and their recovery after THR?

The number of expectations was recorded as a continuous variable indicating the raw number of expectations given by each patient. To investigate the relationship between the number of expectations and recovery, after controlling for potentially confounding clinical and demographic variables, a series of hierarchical multiple linear regression analyses were conducted with each recovery variable as the outcome. Analyses for 'need' and 'desire' expectations were performed separately. Clinical and demographic variables were entered simultaneously as a block first, followed by the number of expectations in the second step.

Research question 3: Is there a relationship between the content of expectations expressed by patients and their recovery after THR? (i.e., are I expectations associated with recovery in a different manner to A expectations or P expectations?)

As described in Section 4.3.3.4, the content of expectations was classified into the ICF constructs of I, A, P or any combination thereof (see Table 4.3 for the number of cases per classification group for each expectation question). For there to be sufficient cases in each classification group to conduct statistical analyses, the number of groups was reduced to five types: I, A, P, AP, and any other combination.

To explore research question 3, a series of multivariate analyses of variance (MANOVA) were conducted. The relationship between the content of 'need' expectations and recovery was investigated separately to the relationship between the content of 'desire' expectations and recovery. Due to the high likelihood of multicollinearity between the Total WOMAC recovery variable and its subscales of pain, stiffness and function, the subscales were not included as dependent variables in these analyses. Recovery variables of Total WOMAC and EQ-5D were the combined dependent variable. One MANOVA was conducted for each of the first three expectations expressed in response to each of the two expectation questions; therefore, six MANOVA were performed in total.

Research question 4: Is there an association between a patient's recovery and having their expectations met?

Exploratory logistic regression analyses were conducted with the binary outcome variable of 'met' vs. 'unmet' expectations, described in section 4.3.3.5. Analyses explored whether recovery measured only by the Total WOMAC scale and the EQ-5D predicted 'met' or 'unmet' expectations group membership; the WOMAC subscales were not included in analyses. The first model assessed whether Total WOMAC recovery and EQ-5D recovery were associated with having 'met' expectations. The second model explored whether recovery continued to predict 'met' or 'unmet' group membership after controlling for pre-operative demographic and clinical factors.

4.3.6 Ethics

Each of the orthopaedic centres involved in the 'EUROHIP' cohort obtained local ethical approval as required.

4.4 Results

4.4.1 Attrition

Of the 1108 patients responding to the expectation questions pre-operatively, 28% did not complete the 1-year follow-up questionnaires. Reasons for attrition are unknown due to some centres being unable to provide this information. To check the representativeness of the follow-up sample, MANOVA analysis compared the pre-operative demographic and clinical characteristics of patients who completed both parts of the study with those who dropped out. A significant main multivariate effect was found (F (11, 622) = 3.33, p <.001). Univariate ANOVA were carried out for each dependent variable in turn using a Bonferroni adjusted value of *p* =.004 to account for multiple tests. Patients remaining in the study cohort at 1-year post surgery had lower WOMAC physical function score (*F* (1, 632) = 8.41, *p* =.004); higher EQ-5D score (*F* (1, 632) = 9.46, *p* =.002); and higher ASA status (*F* (1, 632) = 8.99, *p* =.003). The final sample of 796 patients, therefore, probably over-represents patients with better pre-operative physical functioning and health-related quality of life, yet with a somewhat worse fitness for surgery status.

4.4.2 Description of Sample Before Surgery

The final pre-operative sample comprised 1108 patients, the majority of whom were in their 60’s or 70’s and female (Table 4.2). Only 23% were still employed prior to surgery, whilst the majority had retired. Fifty percent of the sample had educational qualifications post-school with 18% having a university education. Over 70% of the sample was overweight with a BMI of 25 and over. Eighty percent were reasonably fit for surgery scoring 1 or 2 on ASA status and less than 1% scoring 4. The majority (96%) of patients had Kellgren & Lawrence scores of 3 or 4 indicating moderate narrowing or complete loss of joint space. Patients’ health-related quality of life varied widely with EQ-5D scores ranging from -.059 through to 1.00, with 18% of patients reporting a health state worse than dead. The majority of patients (87%) had a Total WOMAC score of 40 or more, generally scoring higher on the subscales of stiffness and physical function than on the subscale of pain.

Table 4.2 Demographic and clinical characteristics of the sample pre-operatively (n=1108)

Characteristic	
Gender	478 M, 629 F
Age (years)	65.1 (11.0)
Employment status (%)	
Employed	22.7
Retired	58.8
Retired early	8.5
Other	9.9
Educational qualifications (%)	
None post school	50.5
College diploma	31.8
University degree	12.9
Postgraduate degree	4.8
BMI	27.6 (4.4)
ASA status (%)	
1 normal, healthy	18.1
2 mild systemic disease	62.1
3 severe systemic disease	18.9
4 life-threatening systemic disease	0.8
Kellgren & Lawrence radiographic grade (%)	
0 no features	0.5
1 minute osteophyte	0.6
2 definite osteophyte	3.2
3 moderate diminution of joint space	48.3
4 no joint space	47.5
Total WOMAC	59.5 (16.0)
WOMAC pain	55.6 (17.8)
WOMAC stiffness	60.8 (20.8)
WOMAC function	60.5 (16.6)
EQ-5D	.40 (.33)

Note. Values are mean (SD), % or years

4.4.3 Research Question 1: Description of Pre-Operative Patient Expectations

The level of agreement between the two researchers regarding the total number and the ICF construct classification of the first three expectations given by each patient for each question was assessed using the Kappa statistic. Kappa values ranged from 873 to .996 which correspond to 'almost perfect' agreement (Landis & Koch, 1977). Discrepancies were discussed until the number of expectations or the code of a given expectation was agreed by consensus. All expectations were coded to the constructs of the ICF (i.e., there were no expectations that could not be classified as I, A or P, or a combination thereof).

Table 4.3 Proportion of expectations in each of the ICF construct classification groups for the first three patient 'need' expectations and the first three 'desire' expectations

	Need expectations			Desire expectations		
	1st	2nd	3rd	1st	2nd	3rd
Impairment (I %)	3.7	4.2	3.2	1.8	1.4	3.0
Activity limitation (A %)	57.8	59.0	58.4	39.5	32.9	36.8
Participation restriction (P %)	5.2	5.5	5.0	9.2	10.6	9.1
Impairment & activity limitation (IA %)	11.5	7.1	7.1	5.0	4.8	3.3
Impairment & participation restriction (IP %)	0.8	0.8	0.5	0.3	0.0	0.0
Activity limitation & participation restriction (AP %)	18.0	21.5	24.4	40.8	47.7	45.6
Impairment & activity limitation & participation restriction (IAP %)	3.1	1.9	1.4	2.6	2.6	2.1
N	1038	731	438	874	568	329

The total number of expectations given by each patient varied widely; the mean number of 'need' expectations was 2.5 (range 0-12) with nine patients responding *"nothing"*; the mean number of 'desire' expectations was 2.4 (range 0-15) with one patient responding *"nothing"*. Table 4.3 reports the proportion of expectations classified to each ICF construct or combination thereof for the first three 'need' and the first three ‘desire’ expectations. The total number of patients included in the classification for each of the 1st, 2nd and 3rd expectations is also provided. The majority of the patient ‘need’ expectations were classified as A; 57.8%, 59.0% and 58.4% of the first, second and third responses given respectively. Most of the patient ‘desire’ expectations were classified as AP; 40.8%, 47.7% and 45.6% of the first, second and third responses respectively. Less than 5% of all expectations were classified as I.

4.4.4 Correlation Analyses

To examine the relationships between study variables, bivariate correlations were computed between the recovery variables (Total WOMAC, WOMAC subscales and EQ-5D); the number of patient expectations, and demographic and clinical variables (see Table 4.4). A better than average recovery indexed by any of the recovery variables (Total WOMAC, WOMAC subscales or EQ-5D) was significantly associated with being younger; lower BMI; fitter ASA health status; and having higher educational qualifications. Gender was not associated with any recovery variable. Reporting a greater number of ‘need’ or ‘desire’ expectations was associated with better recovery on the Total WOMAC and WOMAC pain subscale, but only reporting more ‘desire’ expectations was associated with better recovery on the WOMAC function subscale and the EQ-5D. Neither the number of ‘need’ nor the number of ‘desire’ expectations was associated with recovery on the WOMAC stiffness subscale. Finally, reporting a greater number of ‘need’ expectations was significantly associated with being female, younger and having a higher BMI. Reporting a greater number of ‘desire’ expectations was similarly significantly associated with being female, and younger, but contrastingly, also associated with having a lower BMI and better ASA status.

Table 4.4 Correlation between demographic and clinical characteristics, patient expectations, and recovery variables

	1	2	3	4	5	6	7	8	9	10	11
Demographic and clinical											
1. Gender[a]	-										
2. Age	-.09*	-									
3. BMI	-.03	-.08*	-								
4. ASA status	-.03	.30***	.14***	-							
5. Educational qualifications[b]	.16***	-.16***	.12**	-.10*	-						
Patient expectation											
6. Number ‘need’ expectations	-.12**	-.13**	.10**	-.03	.05	-					
7. Number ‘desire’ expectations	-.08*	-.13**	-.10*	-.10*	.06	.41***	-				
Recovery											
8. Total WOMAC	.00	-.12**	-.12**	-.18***	.20***	.08*	.10*	-			
9. WOMAC pain	.01	-.09*	-.12**	-.18***	.20***	.09*	.10*	.90***	-		
10. WOMAC stiffness	.03	-.09**	-.11**	-.13**	.19***	.04	.07	81***	.75***	-	
11. WOMAC physical function	.00	-.11**	-.10**	-.17***	.19***	.07	.11**	.99***	.83***	.74***	-
12. EQ-5D	.04	-.10**	-.11**	-.11**	.15***	.05	.10*	.63***	.62***	.53***	.60***

Note. [a]Coded as female= 0, male = 1. [b]Coded as none post-school = 1, college diploma = 2, university degree = 3, postgraduate degree = 4. **p* < .05. ***p* < .01. ****p* < .001.

4.4.5 Research Question 2: Number of Expectations Analyses

4.4.5.1 Predicting recovery from disability (WOMAC) by the number of ‘need’ expectations

As presented in Table 4.5, after controlling for demographic and clinical factors the number of ‘need’ expectations did not significantly predict recovery as indicated by the Total WOMAC (*p* = .12), the WOMAC stiffness subscale (*p* = .42) or the WOMAC physical function subscale (*p* = .14). However, the number of ‘need’ expectations *did* significantly predict recovery on the WOMAC pain subscale after controlling for demographic and clinical factors (*p* = .04); patients reporting more ‘need’ expectations had better pain-related recovery than patients reporting less ‘need’ expectations. Each of the final regression models were significant, but only explaining ≤ 8% of the variance in recovery in all cases.

Table 4.5 Summary of results of multiple linear regressions of Total WOMAC recovery and WOMAC subscales on demographic and clinical characteristics and the number of ‘need’ expectations

	Total WOMAC		ΔR^2	WOMAC Pain		ΔR^2	WOMAC Stiffness		ΔR^2	WOMAC Function		ΔR^2
	β (step 1)	β (step 2)		β (step 1)	β (step 2)		β (step 1)	β (step 2)		β (step 1)	β (step 2)	
Step 1												
Demographic & clinical			.07***			.07***			.05***			.07***
Gender[a]	-.04	-.03		-.03	-.02		-.01	-.01		-.04	-.03	
Age	-.06	-.05		-.03	-.02		-.05	-.05		-.05	-.04	
BMI	-.08*	-.09*		-.08*	-.09*		-.08*	-.08*		-.06	-.07	
ASA status	-.14**	-.13**		-.15***	-.15***		-.09*	-.09*		-.13**	-.13**	
Education[b]	.17***	.17***		.17***	.17***		.16***	.16***		.17***	.17***	
Step 2												
Expectations			.00			.01*			.00			.00
‘Need’ expectation		.06			.08*			.03			.06	
R^2 (final model)		.08			.08			.05			.07	
F (for R^2)		8.46***			8.46***			5.84***			7.60***	

Note. [a]Coded as female= 0, male = 1. [b]Coded as none post-school = 1, college diploma = 2, university degree = 3, postgraduate degree = 4. *$p < .05$. **$p < .01$. ***$p < .001$.

4.4.5.2 Predicting recovery from health-related quality of life (EQ-5D) by the number of 'need' expectations

After controlling for demographic and clinical factors, the number of 'need' expectations did not significantly predict EQ-5D recovery (p = .23) (see Table 4.6). The final regression model was significant and explained 4% of the variance in recovery.

Table 4.6 Summary of results of multiple linear regression of EQ-5D recovery on demographic and clinical characteristics and the number of 'need' expectations

	EQ-5D		ΔR^2
	β (step 1)	β (step 2)	
Step 1			
Demographic & clinical			.04***
Gender[a]	.01	.02	
Age	-.07	-.06	
BMI	-.09*	-.09*	
ASA status	-.06	-.06	
Education[b]	.12**	.12**	
Step 2			
Expectations			.00
'Need' expectation		.04	
R^2 (final model)		.04	
F (for R^2)		7.60***	

Note. [a]Coded as female= 0, male = 1. [b]Coded as none post-school = 1, college diploma = 2, university degree = 3, postgraduate degree = 4. *p < .05. **p < .01. ***p < .001.

4.4.5.3 Predicting recovery from disability (WOMAC) by the number of ‘desire’ expectations

After controlling for demographic and clinical factors, the number of ‘desire’ expectations did not significantly predict recovery as indicated by the Total WOMAC ($p = .138$) or any of the three WOMAC subscales, namely WOMAC pain ($p = .112$), WOMAC stiffness ($p = .379$) and WOMAC physical function ($p = .091$) (see Table 4.7[1]). Each of the final regression models were significant explaining ≤ 8% of the variance in recovery in all cases.

---

[1] The beta values for some demographic and clinical variables reported in Table 4.5 and 4.6 are not identical to those reported in Table 4.7 and 4.8. This is because only participants responding to the ‘need’ expectation question were included in analyses relating to it (i.e., Tables 4.5 and 4.6) and similarly, only participants responding to the ‘desire’ expectation question were included in analyses relating to it (i.e. Tables 4.7 and 4.8).

Table 4.7 Summary of results of multiple linear regressions of Total WOMAC recovery and WOMAC subscales on demographic and clinical characteristics and the number of 'desire' expectations

	Total WOMAC		ΔR^2	WOMAC Pain		ΔR^2	WOMAC Stiffness		ΔR^2	WOMAC Function		ΔR^2
	β (step 1)	β (step 2)		β (step 1)	β (step 2)		β (step 1)	β (step 2)		β (step 1)	β (step 2)	
Step 1												
Demographic & clinical			.07***			.07***			.05***			.07***
Gender[a]	-.04	-.04		-.03	-.02		-.01	-.01		-.04	-.03	
Age	-.06	-.05		-.03	-.02		-.05	-.05		-.05	-.04	
BMI	-.08	-.08		-.08	-.07		-.08	-.08		-.06	-.06	
ASA status	-.14**	-.13**		-.15**	-.14**		-.09	-.08		-.13**	-.13**	
Education[b]	.17***	.17***		.17***	.17***		.16***	.16***		.17***	.17***	
Step 2												
Expectations			.00			.00			.00			.01
'Need' expectation		.06			.07			.04			.07	
R^2 (final model)		.08			.08			.06			.07	
F (for R^2)		7.57***			7.52***			5.25***			6.96***	

Note. [a]Coded as female= 0, male = 1. [b]Coded as none post-school = 1, college diploma = 2, university degree = 3, postgraduate degree = 4. *p < .05. **p < .01. ***p < .001.

4.4.5.4 Predicting recovery from heath-related quality of life (EQ-5D) by the number of 'desire' expectations

As detailed in Table 4.8, after controlling for demographic and clinical factors, the number of 'desire' expectations did not significantly predict EQ-5D recovery ($p = .11$). The final regression model was significant and explained 5% of the variance in recovery.

Table 4.8 Summary of results of multiple linear regression of EQ-5D recovery on demographic and clinical characteristics and the number of 'desire' expectations

	EQ-5D		ΔR^2
	β (step 1)	β (step 2)	
Step 1			
Demographic & clinical			.04***
Gender[a]	.01	.02	
Age	-.07	-.06	
BMI	-.09*	-.08	
ASA status	-.06	-.06	
Education[b]	.12**	.12**	
Step 2			
Expectations			.01
'Need' expectation		.07	
R^2 (final model)		.05	
F (for R^2)		4.32***	

Note. [a]Coded as female= 0, male = 1. [b]Coded as none post-school = 1, college diploma = 2, university degree = 3, postgraduate degree = 4. *$p < .05$. **$p < .01$. ***$p < .001$.

4.4.5.5 Post-hoc analysis of the relationship between recovery and responding to the expectation questions

To test whether simply responding to the expectation questions was associated with better recovery, responders were compared with non-responders on recovery indexed by each of the recovery variables; no differences were found (F (5, 806) = 1.04, $p = .39$).

4.4.6 Research Question 3: Content of Expectations Analyses

4.4.6.1 Relationship between the content of 'need' expectations and recovery

Table 4.9 presents descriptive statistics for recovery variables by the ICF content of their first three 'need' expectations. To examine the statistical significance of the identified differences in recovery scores as determined by the ICF content of patients' 'need' expectation, a one-way MANOVA for each of the three 'need' expectations were conducted with the dependent variables of Total WOMAC and EQ-5D recovery. No significant difference was found on the combined dependent recovery variable for patients with differently ICF classified 'need' expectation 1 (F (8, 1326) = 1.03, p = .41), expectation 2 (F (8, 952) = 1.29, p = .25) or expectation 3 (F (8, 566) = .37, p = .94).

4.4.6.2 Relationship between the content of 'desire' expectations and recovery

Table 4.10 presents descriptive statistics for recovery variables by the ICF content of their first three 'desire' expectations. The same multivariate analytic approach used with 'need' expectations was used to investigate the significance of identified differences in recovery scores by the ICF content of 'desire' expectations. Three one-way MANOVA were conducted with the dependent variables of Total WOMAC and EQ-5D recovery; one MANOVA for each of the three 'desire' expectations. No significant difference was found on the combined dependent recovery variable for patients with differently classified 'desire' expectation 1 (F (8, 1146) = .22 p = .99), expectation 2 (F (8, 760) = .81, p = .60) or expectation 3 (F (8, 442) = .74, p = .66).

Table 4.9 Descriptive statistics for recovery variables by the ICF content of the first three 'need' expectations given by patients.

	NEED Expectation 1					NEED Expectation 2					NEED Expectation 3				
	I	A	P	AP	Other	I	A	P	AP	Other	I	A	P	AP	Other
Total WOMAC															
Mean	.04	-.09	-.03	.20	.09	.48	.03	.13	-.03	.07	.54	.06	.07	.16	-.07
SD	.79	1.08	1.09	.91	.91	.79	1.01	.79	.96	.99	.57	1.01	.94	.84	.97
N	24	409	30	107	107	17	302	20	107	48	6	171	14	81	26
WOMAC pain															
Mean	.00	-.08	-.02	.20	.09	.48	.03	.13	-.03	.07	.54	.06	.07	.16	-.07
SD	.92	1.09	1.09	.83	.91	.79	1.01	.79	.96	.99	.57	1.01	.94	.84	.97
N	25	418	30	121	107	17	302	20	107	48	6	171	14	81	26
WOMAC stiffness															
Mean	.03	-.06	.04	.18	.04	.10	.04	.14	-.02	.09	.46	.06	-.27	.11	-.15
SD	.97	1.04	.99	.85	.96	.98	1.00	.88	.94	.93	.66	.97	.95	.93	1.14
N	25	428	30	125	111	18	314	23	109	50	6	177	15	84	27
WOMAC function															
Mean	-.01	-.09	-.02	.19	.10	.54	.05	-.03	-.03	.07	.51	.07	.10	.13	-.02
SD	.79	1.07	1.06	.88	.91	.78	1.01	.90	.94	.98	.57	1.02	.92	.85	.91
N	25	425	30	124	113	17	308	23	110	51	6	178	15	82	26
EQ-5D															
Mean	-.11	-.04	-.04	.11	.02	.23	-.01	-.04	-.09	.21	.43	.02	.12	.06	-.04
SD	1.01	1.02	1.12	.91	1.07	.65	1.03	1.12	1.05	.84	.61	.97	1.02	1.01	1.10
N	27	421	30	120	110	16	311	30	110	48	6	178	14	80	27

Note. I = impairment, A = activity limitation, P = participation restriction, AP = activity limitation & participation restriction, Other = any other combination. Scores are transformed where necessary so that a positive score indicates better than average recovery and a negative score indicates worse than average recovery.

Table 4.10 Descriptive statistics for recovery variables by the ICF content of the first three ‘desire’ expectations given by patients.

	DESIRE Expectation 1					DESIRE Expectation 2					DESIRE Expectation 3				
	I	A	P	AP	Other	I	A	P	AP	Other	I	A	P	AP	Other
Total WOMAC															
Mean	-.18	.00	.12	.07	.12	.30	-.02	.06	.18	-.12	.64	.02	.26	.09	.21
SD	.89	1.04	.98	.93	.88	.61	1.09	1.03	.84	1.00	.38	1.04	.82	.85	.85
N	9	235	50	253	49	4	126	36	199	28	8	82	23	104	13
WOMAC pain															
Mean	-.26	.05	.07	.09	.03	.63	-.01	.13	.19	-.02	.48	.00	.23	.13	.20
SD	.99	1.03	1.04	.86	1.03	.33	1.03	.95	.81	1.02	.54	.96	.84	.83	.82
N	9	242	51	258	50	4	128	39	201	30	8	82	23	108	13
WOMAC stiffness															
Mean	-.32	.02	.07	.04	.06	.05	.01	.13	.11	-.16	.53	-.07	-.05	.10	.47
SD	.98	1.00	.94	.95	1.04	.83	1.6	.92	.90	1.01	.38	1.07	.96	.84	1.00
N	9	244	54	261	50	4	131	39	20	29	8	84	24	108	13
WOMAC function															
Mean	-.14	.00	.17	.05	.16	.21	.01	-.05	.17	-.06	.66	.07	.32	.10	.16
SD	.83	1.05	.97	.94	.84	.68	1.08	1.05	.86	.96	.38	1.05	.80	.87	.82
N	9	242	56	259	51	4	132	39	201	30	8	86	24	107	13
EQ-5D															
Mean	-.09	-.01	.12	.02	.09	.46	.04	-.08	.12	.15	.46	.11	.18	.08	.38
SD	1.02	1.04	.82	.99	1.11	.74	.94	1.05	.89	1.08	.56	.85	.64	.91	.67
N	9	238	57	258	51	4	132	41	201	31	8	85	23	110	14

Note. I = impairment, A = activity limitation, P = participation restriction, AP = activity limitation & participation restriction, Other = any other combination. Scores are transformed where necessary so that a positive score indicates better than average recovery and a negative score indicates worse than average recovery.

4.4.7 Research Question 4: Association Between Recovery and Having Expectations Met

Six hundred and ninety patients (87% of post-surgery sample) responded to the current function question asked 1-year post-surgery: *"What things CAN you do now that you could not do a year ago, as a result of your hip operation?"* Forty-one patients responded *"nothing"* and were therefore assigned to the 'unmet' expectation group (see section 4.3.3.5 for the coding protocol). All of the remaining responses were classified to the constructs of the ICF, i.e., there was no current function that could not be classified as I, A, P, or a combination thereof. The level of agreement between the two researchers regarding the ICF construct classification of the first three current function responses given by each patient was assessed using the Kappa statistic. Kappa values ranged from .708 to .956, which corresponds to 'substantial' and 'almost perfect' agreement (Landis & Koch, 1977). Discrepancies in assigned codes were discussed until agreed by consensus. Table 4.11 presents the relative percentage of responses classified to each ICF construct category. The majority of current function responses were classified as the A construct; 58.1%, 59.5% and 60.2% of the first, second and third responses respectively.

Table 4.11 Proportion of expectations in each of the ICF construct classification groups for the first three 1-year post-operative current patient function responses

	Current function responses		
	1st	2nd	3rd
Impairment (I %)	6.9	5.1	6.6
Activity limitation (A %)	58.1	59.5	60.2
Participation restriction (P %)	1.5	2.7	3.6
Impairment & activity limitation (IA %)	10.3	6.7	3.6
Impairment & participation restriction (IP %)	0.3	0.7	0.0
Activity limitation & participation restriction (AP %)	18.8	23.6	23.0
Impairment & activity limitation & participation restriction (IAP %)	4.0	1.7	2.9
N	649	415	247

Four hundred and seventy-four patients comprised the expectation ‘met’ group and 216 patients comprised the expectation ‘unmet’ group. Assignment to the ‘met’ or ‘unmet’ group formed the binary outcome variable in logistic regression analyses to investigate whether Total WOMAC and EQ-5D recovery was associated with having expectations met. The first model testing the two recovery variables as predictors was able to distinguish between patients with ‘met’ and patients with ‘unmet’ expectations (χ^2 (2, N = 641) = 20.6, p <.001). As presented in Table 4.12, only the EQ-5D recovery variable was a significant predictor in the model with an odds ratio of 1.29 indicating that for every one point increase in EQ-5D recovery score, patients are 1.29 times more likely to have had their expectations met.

Table 4.12 Binary logistic regression predicting likelihood of having pre-operative expectations 'met' by Total WOMAC recovery and EQ-5D recovery – Model 1

	Odds Ratio	95% confidence intervals	*p*
Total WOMAC recovery	1.20	.96-1.45	.102
EQ-5D recovery	1.29	1.03-1.62	.029

The second model explored whether EQ-5D recovery continued to be a significant predictor of 'met' or 'unmet' expectations after controlling for pre-operative demographic and clinical variables. Like model 1, this model was able to distinguish between patients with 'met' and patients with 'unmet' expectations (χ^2 (10, N = 512) = 26.0, $p = .004$). Two variables made a significant contribution to the model. First, the difference between having no post-school education and a university education significantly predicted having expectations met; patients with a university education were 1.99 times more likely to be coded as having their expectations met than those without any post-school qualifications. Second, after controlling for all other variables in the model, EQ-5D recovery significantly predicted having expectations met; for every one point increase in EQ-5D recovery score indicating better recovery, patients were 1.38 times more likely to have had their expectations met (Table 4.13).

Table 4.13 Binary logistic regression predicting likelihood of having pre-operative expectations 'met' by demographic and clinical variables and EQ-5D recovery- Model 2

	Odds Ratio	95% confidence intervals	*p*
Gender	.87	.59-1.29	.498
Age	.99	.97-1.01	.243
BMI	1.03	.98 -1.08	.316
ASA status - 2[a]	.98	.55-1.74	.430
ASA status - 3 [a]	1.54	.74-3.18	.249
Education - college[b]	1.26	.82-1.94	.294
Education – University[b]	1.99	1.03-3.86	.042
Education - postgraduate[b]	2.53	.80-7.97	.113
EQ-5D recovery	1.38	1.14-1.68	.001

Note [a]Reference group is ASA status 1; [b]Reference group is No post-school education.

4.4.8 Post-Hoc Power Analysis

Post-hoc power analysis calculations suggest that the sample size was not adequate; the number of expectations variable was severely underpowered to find an effect on Total WOMAC recovery (power 18.3%). This cohort study was designed by the 'EUROHIP' collaboration primarily to identify the clinical predictors of recovery after THR and power calculations were based on these clinical variables. The role of patient expectations had such a small effect on recovery that the study would have need to have been more highly powered with a larger sample size to have improved the likelihood of finding an effect.

4.5 Discussion

This study has shown that the ICF can be used to classify patient expectations. All expectations could be classified to one or a combination of the ICF constructs, namely I, A and P. Moreover, all responses to the current function question asked 1-year post-surgery, could be classified to the constructs. This supports the use of the ICF as a theoretical framework within which the role of patient expectations and beliefs in recovery after surgery and other interventions can be studied.

The majority of ‘need’ expectations were classified as A and the majority of ‘desire’ expectations were classified as AP. The ICF definitions state that A focuses on the execution of activities whilst P emphasises the individual’s social participation and their involvement in life situations (World Health Organization, 2001).The finding that both ‘need’ and ‘desire’ expectations are related to the execution of activities, but only ‘desire’ expectations are also related to the individual’s social participation, suggests that what patients’ *need* to be able to do and what they *would like* to be able to do may be different. For example, they may need to do basic activities such as walking and bending down but they may also like to be able to participate in social activities such as playing with grandchildren or sport.

Less than 5% of patient expectations were classified to the I construct. This finding conflicts with the body of evidence within osteoarthritis research and practice showing that pain is a measure of I (Brockow et al., 2004; Dreinhofer et al., 2004; Pollard et al., 2006) and pain is a primary symptom of osteoarthritis (Dekker et al., 1992). Further, Mahomed et al’s (2002) study of the relationship between patient expectations and outcomes after joint replacement found that over 75% of patients expected complete pain relief. There are several possible explanations for the discordance between the current finding and previous research. First, pain expectations may not be spontaneously reported by patients. For example, Mahomed et al (2002) asked patients to report the level of pain they expected, whereas the current study did not prescribe that patients respond about their expected pain relief but rather provided them with a space to respond to the expectation questions in any way they chose. Even though impairment-related improvements such as pain relief may be expected by many patients due to the fact that THR targets impairment by removing damaged tissue and reducing pain, pain expectations may only be expressed when directly asked about. Dixon & Johnston (Dixon & Johnston, 2008) similarly found that pain-related cognitions are not spontaneously reported by people with osteoarthritis in the community. Thus, it would seem that even though pain is a primary symptom of osteoarthritis, pain-related cognitions are not necessarily spontaneously expressed by patients.

An alternative potential explanation for the lack of expectations classified as I may relate to the wording of the expectation questions in the current study and the effect of demand characteristics. The questions ask the patient to report what they cannot *do* but need or would like to be able to *do*; hence, the repeated use of the verb ‘do’ may have shaped patients’ thinking towards actions and tasks, and resulted in them reporting A and P expectations. This may explain the finding that the majority of expectations were classified as A and P. However, even if the wording of the questions had primed patients towards giving A and P expectations, if they had still held strong I expectations, then you might have expected to see more expectations classified as IA (e.g., walk without pain) or IP (e.g., watch a film in the cinema with my grandchildren without being in pain), which was not the case. Future work should ask open-ended questions about what patients expect from surgery without using the verb ‘do’ study, which will allow further investigation of the content of spontaneously reported expectations and how question wording may shape responses.

Overall, this study found limited evidence for the role of patient expectations on recovery from THR. The number of ‘need’ patient expectations was found to be positively associated with recovery from pain-related disability (WOMAC pain subscale), after controlling for demographic and clinical variables; the more ‘need’ expectations a patient reported pre-operatively the better their pain-related recovery. However, no other significant relationships were identified between the number of ‘need’ expectations and any of the other recovery variables, or between the number of ‘desire’ expectations and any recovery variable. Moreover, no difference was found between patients expressing different types of expectations on any of the recovery variables; having an I expectation (e.g., pain relief), an A expectation (e.g., walking), a P expectation (e.g., socialising), or an expectation related to a combination of the ICF constructs, did not affect recovery.

Previous research identifying a relationship between patient expectations and pain-related recovery found that the expectation of complete pain relief after joint replacement was an independent predictor of 6-month post-surgery function and improvement in pain levels (Mahomed et al., 2002). Even though the current study found that reporting more ‘need’ expectations was related to a better pain-related recovery, in contrast to Mahomed et al’s (2002) study, we did not find that patients holding an impairment expectation (i.e., pain) faired any better on any recovery variable than patients holding any other type of expectation. It is possible that a patient reporting many expectations is also likely to hold an expectation related to pain simply due to the fact that they have more expectations. Nevertheless, this explanation cannot be investigated because the current study hypotheses tested the independent influence of the content of the first three expectations rather than the influence of holding a certain type of expectation or not.

Potential mechanisms to explain the relationship between patient expectations and surgical recovery have been discussed in the literature. For example, Flood et al (1993) propose mechanisms by which expectations can act to affect outcomes including motivating the patient to cooperate with treatment and improve coping; altering symptom perception; changing the patient's understanding of the disease and guiding information gathering and; altering the patient's anxiety to heighten or reduce symptoms. The current study found that reporting more 'need' expectations was significantly associated with better pain-related recovery; therefore, it is possible that patients who have more expectations about what they expect and need to be able to do, perceive themselves as having more roles and responsibilities. This may mean that these patients are more motivated to fulfil these roles and responsibilities and use more adaptive coping, which in turn may affect their perception of pain and result in better pain-related recovery. However, even if one or a combination of these mechanisms can account for the observed relationship between the number of 'need expectations and pain-related recovery, it is emphasised that the significant finding exists amidst multiple null findings regarding the role of patient expectations on recovery; thus, replication is needed to verify the robustness of the finding.

The last research question addressed whether there is an association between a patient's recovery and having their pre-operative expectations met. Recovery indexed by EQ-5D was found to predict the likelihood of having expectations met after controlling for demographic and clinical variables; for every one point increase in EQ-5D recovery score, patients were 1.38 times more likely to have met expectations. This result indicates that better recovery from general health-related quality of life is associated with THR patients having met expectations. The variable of 'met' or 'unmet' expectations was computed based on whether the content of pre-operative expectations matched the content of responses to a question about current function asked 1-year after THR; therefore, patients with better health-related quality of life recovery may have held pre-operative expectations that were more realistic and thus, more likely to have matched their response to the current function question, resulting in being coded as having ‘met’ expectations. Conversely, patients with worse recovery may have held less realistic pre-operative expectations which did not match their current function resulting in being coded as having ‘unmet’ expectations.

Patient satisfaction is increasingly recognised as an important healthcare outcome. For example, as part of a more patient-centred vision for the National Health Service in the UK, it has recently been decided that up to 10% of a healthcare trust’s income will be dependent on patient satisfaction (Department of Health, 2009). Patient satisfaction may help to explain the current study’s finding. Patient satisfaction has been associated with improved outcomes and recovery in arthritis patients (Ross et al., 1990) and there is some evidence for its association with patient expectations; for example, Mancuso et al (1997) found that satisfaction and expectations were strongly related in 91% of THR patients. Patient satisfaction was not directly assessed in the current study but it is possible that patients with met expectations were also the most satisfied with surgery which may explain the relationship with better recovery. However, a recent study found that even though 93% of TKR patients reported being satisfied with the outcome, their expectations had not been fulfilled, suggesting that satisfaction is not synonymous to met expectations (Nilsdotter, Toksvig-Larsen, & Roos, 2009).

It could also be argued that both better recovery and having met expectations represent indices of good outcome, therefore, the current study’s finding that the two are associated is circular, i.e., a better recovery predicts having met expectations, and having met expectations predicts better recovery. Perhaps, investigating whether pre-operative disability and health status is associated with having expectations met, rather than whether recovery is associated with having met expectations, would address the issue of circularity. However, the main aim of this study was to investigate the relationship between patient expectations and *recovery* after THR. It is also possible that a reciprocal relationship exists between recovery and having met expectations because they were both self-reported, i.e., if I perceive my health-related quality of life to be better, then maybe I feel that my expectations have been met; and if I believe my expectations of surgery have been met, then I may consider myself ‘recovered’, shaping my perception of my health-related quality of life. Current work is exploring the possibility of a biological marker for osteoarthritis before and after joint replacement (Deberg, Dubuc, Labasse et al., 2008), which could be used as an objective measure of health outcome to allow this issue of reciprocity between recovery and having met expectations to be further examined.

Aside from the issues surrounding the possible nature of the relationship between recovery and having met expectations, it is worth highlighting that recovery indexed by the WOMAC was not found to be related to having met expectations. Generic health-related quality of life instruments, such as the EQ-5D, measure general health status whereas the WOMAC measures disability specific to osteoarthritis (Hawker et al., 1995). Thus, it seems that it is a patient’s general health-related recovery rather than disease-specific recovery which characterises the identified relationship between recovery and expectations. Replication of the significant relationship between better EQ-5D recovery and having met expectations and the lack of significant finding for WOMAC recovery would be useful to test the robustness of findings. The possibility that met expectations is synonymous to patient satisfaction, and that met expectations can account for the variation in patients' health-related quality of life recovery could also be further explored.

This study found limited evidence for the role of patient expectations on surgical recovery. Evidence from other studies has been mixed. Some studies have found that positive expectations are related to better health outcomes after surgery, after controlling for demographic and clinical factors (de Groot et al., 1999; Henn et al., 2007; Lutz et al., 1999; Mondloch et al., 2001; Orbell et al., 1998), whereas others have not (Flood et al., 1993; Mannion et al., 2009). A potential explanation for the mixed findings that exist in the literature is the lack of application of a theoretical framework within which the role of patient expectations on recovery can be studied. Moreover, variation exists between studies in the measurement and conceptualisation of patient expectations. The current study elicited expectations via a free-response question allowing the content of patient expectations to be examined with reference to a theoretical framework, namely the ICF. Classifying patient expectations according to the ICF in future studies will allow direct comparisons between studies, thereby supporting the development of a cumulative evidence base.

The main limitation of this study relates to the wording of the expectation questions. As previously discussed, the questions repeated use of the verb 'do' may have encouraged A and P expectations to be reported rather than I expectations. Further, in addition to the semantic difference between the expectation questions by the use of *"need"* or *"would like"* to elicit needs or desires respectively, there is another difference in the wording of the questions. The 'need' item asks patients to identify what they think they will be able to do and need to be able to do, whereas the 'desire' item only asks them to identify what they would like to be able to do. Thus, it could be argued that only the 'need' item elicits an expectation and the ‘desire’ item simply elicits a desire or wish. The evidence indicates that patients’ needs are different to their desires; the majority of ‘need’ expectations were classified as A whilst the majority of ‘desire’ expectations were classified as A and P. Furthermore, needs may have a different relationship with recovery than desires; the number of ‘need’ but not ‘desire’ expectations was associated with pain-related recovery. Careful consideration of question wording is needed to elicit patient expectations in future work.

There are other limitations of this study that should be acknowledged. First, the follow-up sample had better pre-operative physical function and health-related quality of life but worse ASA status; therefore, it is possible that the sample does not fully represent recovery after THR, but is only applicable to patients who were less debilitated by the disease pre-operatively. Second, this study utilised the ‘EUROHIP’ cohort, which is essentially a convenience sample of THR patients recruited from orthopaedic centres with an interest in the prospective investigation of THR patients. Therefore, it is possible that the participating centres differ in some important way to non-participating centres. Similarly, it is possible that patients who agreed to participate differ on some demographic, clinical or psychological variable to those who refused to take part, potentially further weakening the representativeness of the sample. Study replication would therefore be useful. Third, a measure of the strength of the expectation was not available but may be an important factor to consider. For example, many patients held A expectations about being able to walk after surgery; perhaps, patients who anticipated being able to walk as *extremely likely* had better recovery than patients who only anticipated being able to walk as *likely*. Future work could employ a free-text response question as was employed in the current study followed by a Likert-type item to assess the strength of the expectations reported.

Finally, this study computed a valid measure of recovery used in previous studies of health outcomes (Johnston et al., 1999; Johnston et al., 2007; Johnston et al., 2004; Molloy et al., 2009). However, other studies have assessed health outcome using post-operative health and functioning scores and controlling for pre-operative scores in the regression analyses (Mahomed et al., 2002). In the current study we opted against this approach because post-operative scores were highly skewed; floor effects were seen for the WOMAC with many patients reporting a lack of disability and ceiling effects were seen for the EQ-5D with many patients reporting the best possible health-related quality of life. Transformation of these variables was not able to normalise the distributions. Possibly less affected by floor and ceiling effects is the use of *difference* or *change scores* between pre-operative and post-operative scores. However, change scores are criticised for the problem of regression to the mean because they tend to be negatively correlated with time 1 scores (Taris, 2000), meaning that patients with higher pre-operative scores are likely to display relatively smaller gains than those with lower pre-operative scores. Therefore, this study used residualised scores of recovery. Unfortunately, even though the recovery variables were more normally distributed than the post-operative scores, they were computed using the post-operative scores which may explain the heteroscedasticity evident in the recovery variables. Heteroscedasticity is said to weaken but not invalidate analyses (Tabachnick & Fidell, 2001), however it does highlight a limitation with the measurement sensitivity of the WOMAC and EQ-5D as health outcome measures after THR, as many patients scores showed floor and ceiling effects respectively.

4.5.1 Summary and Implications

This study has shown that the ICF is a suitable framework to classify patient expectations and explore their relationship with recovery after THR. Application of the ICF could be extended to investigate the role of patient expectations and beliefs on recovery after other types of surgery and interventions. Limited evidence for the role of patient expectations on recovery after THR was found. However, after controlling for demographic and clinical variables, reporting more ‘need’ expectations was associated with better pain-related recovery and; better recovery from health-related quality of life was associated with having ‘met’ expectations. However, the questionnaire items used to obtain patient expectations may have elicited behavioural expectations at the expense of impairment-based expectations. Future studies should employ free-text response questions that ask patients to express their expectations in a manner that is free from any such demand characteristics.

Patient cognitions are amenable to change. For example, illness perceptions can be modified by intervention, improving functional outcomes in myocardial function patients (Petrie, Cameron, Ellis et al., 2002). Therefore, it follows that patient expectations may also be modified for optimal recovery. For example, patient expectations could be elicited by clinicians during a pre-operative assessment and encouragement could be given so that patients formulate more ‘need’ expectations to promote better pain-related recovery. Similarly, ensuring that patients hold realistic pre-operative expectations may be key in determining that a patient’s expectations are met, which may result in better health-related quality of life recovery.

Chapter 5 - The *n-of-1* methodology and experimental design

Chapters 6 and 7 describe a series of behaviour change intervention single case studies. Both chapters apply the same *n-of-1* methodology and experimental design to investigate the behaviour of physical activity (PA). However, individuals from different sample populations are studied in each chapter. Chapter 6 investigates PA (disability) behaviours in individuals with mobility problems associated with osteoarthritis and Chapter 7 investigates PA behaviours in healthy individuals for comparison. The Theory of Planned Behaviour (TPB) is the theory applied in these behaviour change studies as it is a component part of the integrated International Classification of Functioning, Disability and Health (ICF) /TPB model applied in this thesis. Individuals from a population with mobility problems and individuals from a healthy population are studied to permit application of the theory and the *n-of-1* experimental methodology in two distinct samples. This chapter will provide the background and context for using this methodology and describe the experimental design used in both studies. The methodology section in Chapters 6 and 7 will therefore be brief, only providing additional information unique to that particular study.

5.1 Background and Rationale

5.1.1 Current Use and Evidence in Support of the Theory of Planned Behaviour

As discussed in Chapter 1 Section 1.3.3, the TPB (Ajzen, 1991) is one of the most extensively applied social cognition models in the study of health behaviours, including PA (Armitage & Conner, 2001; Hagger et al., 2002; Sheeran, 2002). The TPB proposes that behavioural intention and perceived behavioural control are the most proximal predictors of subsequent behaviour. While the concept of intention is said to capture an individual’s motivation to perform the behaviour, perceived behavioural control is said to capture an individual’s perceptions of control over performing the behaviour. Perceived behavioural control is similar to Bandura’s (1997) construct of self-efficacy, which refers to a person’s confidence in their ability to perform the behaviour. It has been widely debated whether perceived behavioural control and self-efficacy are distinct or interchangeable constructs (Ajzen, 1991; Bandura, 1992b; Terry & O'Leary, 1995), or rather whether perceived behavioural control is a concept that refers to both controllability and self-efficacy beliefs (Ajzen, 2002); operationalisation of a heterogeneous perceived behavioural control construct would require measures of each subcomponent, i.e., perceived controllability and self-efficacy (Ajzen, 2002).

There is compelling evidence for the predictive value of the TPB towards a variety of behaviours (Armitage & Conner, 2001; Godin & Kok, 1996; Hagger et al., 2002). However, evidence addressing the causal structure of the theory by its application in behaviour change interventions is somewhat scarce (Hardeman, Johnston, Johnston et al., 2002). More experimental investigation of the TPB is needed. The application of behaviour change interventions based on the TPB will permit the causal structure of the theory to be tested by targeting one or more of the proposed causal determinants of behaviour, such as intention or perceived behavioural control, and observing the effects on behaviour.

Webb & Sheeran (2006) conducted a meta-analysis of experimental studies targeting intention (within a variety of theoretical frameworks not exclusively those based on the TPB) and concluded that medium-to-large changes in intention lead to small-to-medium changes in behaviour. Interventions identified as more likely to be successful in generating intention and behaviour change did however include those based on the TPB.

A randomised controlled trial by Kelley and Abraham (2004) targeted intention and perceived behavioural control to increase healthy eating and PA and concluded that the intervention successfully produced higher gains in the targeted cognitions and behaviour in the intervention group than in the control group. Other studies have shown that interventions targeting constructs of perceived controllability and self-efficacy can produce changes in the targeted cognitions and in behaviour, indicating that the change in behaviour is mediated by the change in cognitions (Fisher & Johnston, 1996; Luszczynska & Tryburcy, 2008). Indeed, there is strong empirical evidence in support of self-efficacy as a casual cognition within the self-efficacy theory literature (Bandura, 1992a). Hence, even though this evidence did not arise from interventions explicitly based on the TPB, experimental manipulation of the targeted cognitions of perceived controllability and self-efficacy, as subcomponents of perceived behavioural control, provides preliminary support for a causal relationship between perceived behavioural control and behaviour as set out in the TPB.

Another intervention commonly associated with the TPB is the formation of implementation intentions or ‘action plans’ (Gollwitzer, 1993). In contrast to interventions targeting the proposed causal determinants of behaviour such as behavioural intention or perceived behavioural control, this intervention employs a volitional strategy where the individual makes a plan specifying *when*, *where* and *how* a goal-directed behaviour will be performed. The plan specifies the situational cues in which the behaviour is to be enacted; thus, connecting good opportunities to act with a behavioural response (Gollwitzer & Sheeran, 2006). In essence, implementation intentions can be regarded as a volitional strategy to transfer behaviour control to the environment. This strategy can aid effective self-regulation of goal striving by facilitating the translation of intentions into behaviour and bridging the intention-behaviour gap (Gollwitzer, 1999; Gollwitzer & Sheeran, 2006).

Sniehotta (2009) has recently proposed a distinction between interventions employing a *planning* approach and those more strictly adhering to the implementation intention paradigm. He argues that the implementation intention paradigm involved controlled stimulus-response experiments where the participant is instructed to perform behaviour *y* when they see the specific cue *x*; whereas *planning* interventions in health psychology have typically relied on the participant completing their own plan and defining the situation and behaviour. Furthermore, Sniehotta (2009) argues that the effect of implementation intentions was determined by the performance of behaviour *y* in response to situation *x*, whereas health psychology *planning* interventions typically involve creating a conditional plan, such as "On Monday at 2pm (*x*), I will swim in the University pool (*y*)", but the behavioural goal (*X)* which is measured to determine the effect of the intervention, is *un*conditional, such as being more physically active. Therefore, in order to avoid the possible further amalgamation of distinct theoretical concepts, the term 'action planning' will be used hereafter to refer to the type of planning intervention described by Sniehotta (2009).

The 'action planning' technique has received much attention across a range of health behaviours. Some studies have concluded that action planning interventions can successfully change behaviour (Gratton, Povey, & Clark-Carter, 2007; Luszczynska, 2006; Kellar & Abraham, 2005; Sheeran & Silverman, 2003), whereas other have been unable to conclude this (Jackson, Lawton, Knapp et al., 2005; Michie, Dormandy, & Marteau, 2004; Rutter, Steadman, & Quine, 2006). Nonetheless, a recent meta-analysis of interventions employing either implementation intentions or action planning, concluded that planning interventions can have a medium-to-large positive effect on the performance of many behaviours (Gollwitzer & Sheeran, 2006).

Despite the ever-increasing number of theory-based intervention studies investigating the potential causal pathways between TPB construct, support for the causal structure of the TPB is extremely limited by the dominance of group-based design studies. Group-based intervention studies examine differences between individuals and interventions target predictors of individual differences. An evidence-based TPB intervention targeting one or more of the potentially causal constructs, such as intention or perceived behavioural control, can further our knowledge of the components and process mechanisms intrinsic to the effectiveness of the intervention (Lippke & Ziegelmann, 2008; Michie & Abraham, 2004; Michie, Johnston, Francis et al., 2008). Yet, in the case of a typical group design study, the data from all participants is pooled for statistical analyses and the effectiveness of an intervention is evaluated by drawing a comparison of the average effects across groups. This means that the group findings cannot be extrapolated to represent the specific effect on any individual within the group (Ottenbacher, 1990) nor can interventions be tailored to target specific predictors for an individual.

While studies frequently find that stronger intentions and higher perceived behavioural control predict more engagement in a behaviour, few studies have investigated whether an individual is more likely to engage in the behaviour at times when they have stronger intentions and higher perceived behavioural control than at other times. Nevertheless, implicit in the theoretical framework of the TPB is the assumption that the theory should apply *within* individuals. In order to address the question whether an individual is more likely to engage in a behaviour when their intention is stronger and perceived behavioural control is higher, within-individual variability in the reported strength of these constructs is required. Variability in TPB cognitions has been reported in studies examining whether the temporal stability of intention and perceived behavioural control moderates the relationship between cognitions and behaviours (Conner, Sheeran, Norman et al., 2000; Sheeran & Abraham, 2003; Sheeran, Orbell, & Trafimow, 1999); however, despite providing some evidence for variability in TPB cognitions over time, the between-person analyses used in these studies preclude further investigation of variability within individuals. In contrast, within-individual variability in condom use intention and self-efficacy has been reported in a study employing a within-person design, which found that day-to-day variability in condom use intentions was associated with failure to use condoms (Kiene, Tennen, & Armeli, 2008).

5.1.2 The n-of-1 Experimental Methodology

An alternative to the group design is the study of individuals with an *n-of-1* (single subject) design. Indeed, the new Medical Research Council Complex Interventions Guidance identifies *n-of-1* methods as an important tool for theory testing (Craig et al., 2008). Inherent to the *n-of-1* design is the potential to examine within-individual variability, to test theory within individuals and to test the effectiveness of an intervention for a specific individual.

The *n-of-1* design has been used in a range of health-related fields including neuropsychological rehabilitation (Robertson, Hogg, & McMillan, 1998), pharmaceutical trials (March, Irwig, Schwarz et al., 1994; Guyatt, Heyting, Jaeschke et al., 1990), physiotherapy treatment (Asenlof, Denison, & Lindberg, 2005; MacDonald, Whitman, Cleland et al., 2006; Cowell & Phillips, 2002) and clinical psychology (Fisher & Wells, 2008; Jones, Johnston, & Speck, 1989). The design involves the repeated measurement of variables in an individual over a period of time, which not only allows variability in the measured constructs to be observed but also the theoretically and clinically important relationships between variables to be determined. For example, use of an *n-of-1* design refuted the hypotheses that depression predicted an increase in symptoms in a patient with irritable bowel syndrome and bipolar disorder over a 12-month period (Crane, Martin, Johnston et al., 2003).

Advocates of the *n-of-1* approach highlight the fact that the method permits investigation of more rare syndromes or behaviours and easily dovetails clinical practice whereby the clinician selects the appropriate patient to participate in the study (Shallice, 1979). As argued by Canavan (Canavan, 1994) the *n-of-1* methodology should not replace group studies but that the two approaches are complimentary. Conclusions from *n-of-1* experimental studies can help develop hypotheses to be tested further in a large-scale trial and hypotheses derived from conclusions from a large scale trial can be empirically tested within individuals (Onghena & Edgington, 2005).

Experimental manipulation of the proposed causal determinants of behaviour change within the TPB is needed in order to test the causal pathways of the theory within individuals. Systematic *n-of-1* trials can apply TPB-based interventions to test whether the theory is supported within individuals and also to provide valuable knowledge regarding effective theory-based behaviour change techniques for individuals. The *n-of-1* methodology involves close and regular monitoring of the variables under investigation. For this reason interventions can be data driven and designed in response to the individual’s data (Morley, 1994). To test the effect of an intervention in a specific individual, several research designs are possible with data obtained in non-intervention periods, acting as control conditions (see Barlow, Nock, & Hersen, 2009 for a review of experimental designs within n-of-1 studies). Interventions can also be tailored to the individual and their setting, and changes as a result of an intervention can easily be evaluated on individually specified outcome measures. There is evidence that personally tailored health behaviour interventions are preferred by the recipient and more effective in promoting health behaviour change when compared to standard interventions (Ryan & Lauver, 2002).

5.1.3 Ecological Momentary Assessment, Diary Methods and the n-of-1 Methodology

The investigation of daily events and experiences are increasingly seen in the health-related and behaviour change literature. Affleck et al (Affleck, Zautra, Tennen et al., 1999) use the term ‘daily process studies’ to include studies with designs that involve the repeated measurement of variables believed to vary from day to day in some meaningful manner; thus each variable can be deemed to be a daily process. Different methods for recording daily events are available and will depend on the key research questions being addressed. Ecological momentary assessment approaches include methods that allow an individual’s behaviours and experiences to be studied in their natural environment and in real time (Shiffman, Stone, & Hufford, 2008). The individual completes a brief assessment of their current behaviour and perhaps their mood and thoughts related to the behaviour in paper or electronic diaries over several days or weeks. Individuals are generally “beeped” by the electronic diary or alerted in some other way to complete the assessment at predetermined regular intervals or via random time sampling (Stone, Kessler, & Haythornthwaite, 1991).

Daily diary methods are an ideal choice for use in studies with an *n-of-1* design; within-individual assessments can easily be collected over time. Assessments at equal intervals will support analyses examining within-individual variability and the relationships between variables, whereby the time interval between diary entries serves as the unit of analysis (Shiffman et al., 2008). The total number of assessments and the frequency with which assessments are made depends on the specific study design. Moreover, an experimental *n-of-1* study needs to ensure that a stable baseline period is obtained prior to the initiation of an intervention in order to detect the true effects of the intervention. The more observations per day and/or a longer study period increases the statistical reliability of the data but is also likely to increase participant burden (Stone & Shiffman, 2002), which may result in poor diary compliance and participant attrition.

Another issue related to diary compliance has been investigated comparing electronic and paper diaries. Stone et al (2002, cited in Stone & Shiffman, 2002) used a paper diary with a photosensor to covertly record opening of the diary and found that participants reported compliance to diary completion 90% of the time, yet actual compliance assessed by the photosensor was only 11%. Participants who were given an electronic diary which prompted assessments revealed timely compliance 94% of the time. A meta-analysis of electronic and paper assessments of patient-reported measures concluded that both methods produce equivalent results in terms of test-retest reliability and mean differences between measures, yet the electronic method is likely to increase compliance (Gwaltney, Shields, & Shiffman, 2008).

The studies reported in Chapters 6 and 7 employ an experimental *n-of-1* design using a daily diary method to test the TPB within individuals. The exemplar health behaviour is PA behaviour.

5.1.4 Physical Activity Behaviour

Physical inactivity has been internationally identified as one of the leading causes of death (Mokdad, Marks, Stroup et al., 2004; World Health Organization, 2003). Further, a sedentary lifestyle has been said to explain nearly a quarter of all preventable deaths in the United Sates (Booth, Gordon, Carlson et al., 2000). There is strong evidence that regular PA reduces the risk of many illnesses and chronic conditions including obesity, heart disease, diabetes, cancer, stroke and musculoskeletal complaints (Department of Health, 2004b; World Health Organization, 2003). The psychological benefits of PA are also well-documented including anxiolytic, antidepressant and stress-reducing effects (Salmon, 2000).

Much of the literature focuses on PA as a primary preventive health behaviour. PA is a modifiable behaviour that can prevent disease and disability and therefore, has the potential to compress morbidity, especially in elderly populations where morbidity is currently highest (Fries, 1996). PA is also an effective secondary preventive health behaviour that can be included in the management of an existing disease or in the prevention of further episodes (Kaplan, 2000). For example, PA is recognised in the management of hypertension, depression, obesity and osteoarthritis (National Institute for Health and Clinical Excellence, 2006; National Institute for Health and Clinical Excellence, 2008). Thus, PA can also be conceptualised as a rehabilitation behaviour and in accordance with the definition of disability as behaviour, PA can also be conceptualised as disability behaviour (activity limitation) in individuals with an underlying disabling health condition such as osteoarthritis.

Public health guidelines recommend that all adults engage in at least 30 minutes of moderately intense aerobic PA on five or more days of the week (Department of Health, 2004a; Haskell, Lee, Pate et al., 2007). This recommendation does not differentiate between younger, middle and older age adults, nor does it differentiate between those who are healthy and those with long-term conditions such as osteoarthritis. Despite the recommendation, the most recent figures report that only 28-35% of women and 40-46% of men in the UK (National Health Service Information Centre for health and social care, 2009; The Scottish Government, 2008); and 49% of adults (men and women combined) in the USA (National Center for Chronic Disease Prevention and Health Promotion: 2007) meet the current recommended levels of PA per week.

Intentional PA is a volitional behaviour. An individual's decision to engage in regular PA may be influenced by one or more actual or perceived barriers. For example, time and cost restraints, the physical environment such as the weather, or the individual's physical ability or attitude towards PA may function to deter the individual. In addition, some forms of PA may require access to facilities or equipment such as a swimming pool or gym, others may simply require another person to play with such as tennis. Exercise-referral schemes exist whereby individuals in primary care are referred to a tailored PA programme held in an exercise facility; however, there has been little evidence that such schemes increase PA (National Institute for Health and Clinical Excellence, 2006; Williams, Hendry, France et al., 2007).

In contrast, walking is a form of PA that is not facility dependent and can be sustained into old age (Morris & Hardman, 1997). It can also be performed free of cost and alone. Walking at a pace of 3 miles an hour expends enough energy to meet the definition of moderately intense activity (Ainsworth, Bassett, Strath et al., 2000) and therefore, walking at this pace for 30 minutes five days a week would fulfil the current recommendations. Interventions promoting walking have been identified as able to produce sustainable increases in PA (Hillsdon & Thorogood, 1996; Ogilvie, Foster, Rothnie et al., 2007). Subsequently, walking behaviour has been the focus of many public health campaigns. People are advised to take 10,000 steps a day to help them achieve the recommended 30 minutes of PA a day; yet, on average people only take between 3,000 and 5,000 steps (British Heart Foundation, 2009). Further evidence has suggested that in order to meet the guidelines, walking intensity should equate to 3,000 steps in 30 minutes (Marshall, Levy, Tudor-Locke et al., 2009).

5.1.5 General Research Questions Addressed in Chapters 6 and 7

Chapters 6 and 7 use an *n-of-1* experimental design to test the ability of the TPB to explain PA behaviours and PA behaviour change within individuals with osteoarthritis (Chapter 6) and within healthy individuals (Chapter 7).

Two key research questions were investigated in each study:

1. Does the TPB predict PA *within* individuals?
2. Do individually tailored interventions increase PA *within* individuals?

5.2 Design

Diary methods were used to observe variability in the proximal predictors of the TPB and PA behaviours studied within individuals over a period of 12 weeks. A non-intervention baseline period allowed predictive and therefore potentially causal relationships between the proximal predictors of the TPB and PA to be identified in each individual. These data directed the type of intervention the individual would receive: (a) a TPB-based intervention designed to increase PA by increasing one or other of the components of perceived behavioural control, namely perceived controllability (PC) or self-efficacy (SE), or (b) an action planning intervention designed to increase PA, by aiding the transition of intentions into behaviour.

5.3 Participants

A series of single case studies in four individuals with osteoarthritis are reported in Chapter 6 and in six healthy individuals are reported in Chapter 7. The rationale behind studying multiple individuals was to maximise the data and the potential to identify more factors that contribute to explaining PA behaviour. Specific information about the participants is provided in the methodology section of the respective chapters.

5.4 Measures

Measures that were common to both studies are reported here. Information relating to the measurement rating scales is reported in each of the respective chapters.

5.4.1 Behaviour

Walking behaviour for each individual was assessed objectively by pedometer step count and recorded by the participant at each diary entry. The MACTAR questionnaire (Tugwell, Bombardier, Buchanan et al., 1987) was used with each participant to identify a PA behaviour they were motivated to do, in addition to the PA behaviours prescribed in the study protocol, namely walking in the study reported in Chapter 6, and walking and gym in the study reported in Chapter 7. This personally identified PA was then defined in terms of its Target, Action, Context and Time (TACT principle: Ajzen, 2006). These personally identified behaviours were assessed by self-report and recorded in the diary.

5.4.2 TPB Cognitions

Intention, PC and SE cognitions, in relation to each investigated behaviour, were assessed by the following single TPB items with verbal anchors: Behavioural Intention: 'To what extent do you intend to X between now and the next time you fill in the diary?' (*no intention – definitely intend*); Perceived Controllability (PC): 'How much control do you have over X between now and the next time you fill in the diary?' (*no control – complete control*); and Self-Efficacy (SE): 'How confident are you that you can X between now and the next time you fill in the diary?' (*not at all confident – extremely confident*). 'X' was replaced with each behaviour. These single items have been previously used in an *n-of-1* study (Schroder, 2008) and were developed in accordance with published guidelines for generating TPB items (Francis, Eccles, Johnston et al., 2004). Single rather than multiple item measures of each TPB construct were used to reduce participant burden associated with the length of the daily diary and to enhance response likelihood.

5.4.3 Health Status Measures

In order to compare participants with the general population, the following measures were taken pre- and post-study: 1) SF-36 questionnaire (Ware, Snow, Kosinski et al., 1993) as a valid measure of perceived health status in the general population (Brazier, Harper, Jones et al., 1992) yielding a score from 0 (worst possible health state) to 100 (best possible health state) for the subscales of physical functioning, social functioning, role physical, role emotional, mental health, vitality, bodily pain, and general health; and 2) the HADS (Zigmond & Snaith, 1983) to screen for anxiety and depression in the community (Dowell & Biran, 1990), whereby scores of 8 to 10 are 'possible' clinical disorders and scores of 11 to 21 are 'probable'.

5.5 Intervention Types

5.5.1 TPB-based Intervention Targeting Either PC or SE

This intervention aimed to increase PA behaviour by increasing one or other of PC or SE. The content of the intervention was based on a successful experimental manipulation of control beliefs (Fisher & Johnston, 1996). The following instructions, using walking behaviour in this example, were given to the individual. Wording was adapted as shown for the PC and *SE* interventions, respectively:

> ‘One of the things that influences whether you as individual walk more than usual, is your sense of control/*confidence* over walking. The more control you believe you have/*confident* you feel, the better you will succeed at walking more than usual. Please tell me about three occasions when you felt in control of/*confident* about walking. It may help you to visualise the occasions.’

5.5.2 Action Planning Intervention

Based on the recommendations by Gollwitzer (1993) (Gollwitzer, 1993) and following previous studies (Milne, Orbell, & Sheeran, 2002; Sniehotta, Scholz, & Schwarzer, 2006), the following instructions, using walking behaviour in this example, were given to the individual:

> ‘One of the things that influences whether you as an individual walk more than usual is your intention. However, many people find that despite intending to walk they don’t always manage to carry out their intention and actually do so. You may have even experienced this yourself. It has been found that if you form a definite plan of exactly when you will walk you are more likely to actually do

so. Thinking in terms of every week, what day of the week, time of the day, place, length of time and, if appropriate, with whom do you intend to walk?’

5.6 Procedure

Each participant attended a pre-study session with the researcher which included the MACTAR questionnaire (Tugwell et al., 1987) to identify their personally defined behaviour, instruction on pedometer use, negotiation of timings to fill in the diary and the opportunity for clarification of any diary items. Participants completed the diary twice daily, once in the morning and once in the evening with approximately 12 hours between entries, for 12 weeks with approximately the first 6 weeks constituting the baseline and the subsequent 6 weeks constituting the post-intervention period. The diary recorded PA behaviours (i.e., pedometer step count and self-report) and TPB cognitions. Twice daily, participants were prompted to fill in the diary and participants were telephoned weekly to troubleshoot any problems. The method of prompting diary completion differed in the two studies; hence specific details will be given in the respective chapters. Contact was made with the participant during week 6 to initiate the intervention and again at the end of week 12 for participant debriefing and an opportunity for them to give feedback on their participation.

5.7 Intervention Protocol

At the end of the baseline data collection, data was analysed for each participant individually to identify which cognition (i.e., intention, PC or SE) was most highly correlated with each PA behaviour (see section 5.8 for further detail). A convenient time was arranged to conduct the intervention, where participants were told they would receive individualised feedback to help them increase performance of a chosen PA behaviour. Each intervention lasted approximately 15 minutes and was either telephone-based or face-to-face. Each individual was asked to choose one PA behaviour, from those they reported in the diary, that they were most motivated to do more frequently or for longer periods.

When PC was the cognition most highly correlated with the PA chosen by the participant, the intervention targeting PC was given. When SE was the cognition most highly correlated, the intervention targeting SE was given. When intention was the cognition most highly correlated with the chosen PA, the action planning intervention was given. This decision was based on the empirical support and conceptualisation of action planning (implementation intentions) as a post-intentional strategy to facilitate the translation of intentions into performance of behaviour (Gollwitzer, 1993; Gollwitzer & Sheeran, 2006). When no cognition could be identified, the action planning intervention was offered as the default intervention. This decision was based on the knowledge that, even though a baseline association between intention and PA had not been identified, the action plan would specify the *when*, *where* and *how* the individual would behave in order reach the higher goal intention of increasing the PA they identified as one which they were motivated to do.

To complete the intervention, participants were given a written task designed to reinforce their engagement with the intervention content. Individuals receiving the action planning intervention were asked to complete a written version of their plan and individuals receiving PC or SE interventions were asked to write descriptions of the three occasions. When the intervention was delivered via the telephone, participants were emailed the task and required to email a copy of the completed task back to the researcher. When the intervention was face-to-face this written task was completed in person.

5.8 Analyses

Data were analysed using time series analyses, which is recommended for within-individual investigation (Hamaker, Dolan, & Molenaar, 2005). In accordance with recommendations for time series data (Morley & Adams, 1991), data were also subjected to graphical analysis via visual inspection. Because daily measures from the same individual were collected sequentially, it is likely that data series exhibit a pattern of autocorrelation (i.e., the closer in time two measures of the same variable are taken, the more similar they will be). Thus, to ensure independence between data points, the following procedure was carried out to test for serial dependency within each series and, where identified, it was controlled for using a pre-whitening technique. Only data series exhibiting autocorrelation that exceeded 95% confidence intervals (CI) were subjected to this procedure. First, autocorrelation functions were computed. Possible cyclical relationships within each data series were assumed to be evident in data over a period of a week; thus, a maximum time lag of one week was adopted (i.e., 14 data points). Second, partial autocorrelation functions, which control for autocorrelation at intervening time points, were examined in each series and the time lag at which the largest autocorrelation was seen was identified. Third, each data series was lagged accordingly (i.e., first-order autoregressive relationships were lagged by one time interval, second-order relationships were lagged by two time intervals etc.). Fourth, the lagged series was then regressed onto the original series and the residuals saved. These residuals formed a pre-whitened data series used for all subsequent analyses.

After autocorrelation had been controlled for, TPB cognition series were cross-correlated with PA behaviour series in order to identify the cognition most highly correlated with each behaviour. The association between each cognition and PA was indicated by correlation coefficient functions at each time lag, and the valence of time lag indicates the order in which the two observations occur (e.g., a positive lag between intention and PA indicates that intention precedes PA, whereas a negative lag indicates that PA precedes intention). As with the pre-whitening technique, a maximum lag of one week (i.e., 14 observations) was adopted; however, only cross-correlations of +/-0.40 that also exceeded 95% CI were considered to ensure that identified correlations were statistically significant and could explain a reasonable proportion of variance (Cohen, 1988). This procedure was applied to data from the baseline period to establish the predictors of PA for each individual. It was then applied to data from the post-intervention period to examine either the relationship between the targeted cognition and PA in the individuals given the PC or SE intervention, or the relationship between intention and PA in the individuals given the action planning intervention.

Intervention effects on PA were examined for each individual comparing pre- and post-intervention data. If the participant received either of the PC or SE intervention, then intervention effects on the targeted cognition were also examined comparing pre- and post-intervention data. If the participant received the action planning intervention, then intervention effects on intention were not examined because action planning is not intended to target intention per se. Depending on the normality of data distribution, intervention effects were assessed using either Mann Whitney *U* tests or unpaired *t* tests. Effect sizes are reported as Cohen's *d* statistic, whereby small, medium and large effects estimated as .20, .50, and .80, respectively (Cohen, 1992).

5.8.1 Criteria to Evaluate Theory

The TPB was the theory applied in these series of single case studies. The theory was evaluated by its application within each distinct individual and not evaluated by its application between individuals. The PC or SE intervention tested a potentially causal pathway in the TPB; therefore, the effect of the intervention contributed to the evaluation of the TPB. The action planning intervention was not hypothesised to increase intention nor directly test a potentially casual pathway in the TPB; therefore, it did not contribute to the evaluation of the theory but did contribute to the overall evaluation of action planning. The following criteria were employed to evaluate the TPB:

- When one or more of the theoretical constructs of intention, PC and SE predicts behaviour in the individual, then the evidence provides support for this component of the theory.

- When none of the constructs predict behaviour in the individual, then the evidence does not provide support for the theory.

- When the PC or SE intervention increases the targeted cognition and an increase in behaviour is observed within the individual, then the evidence provides support for the causal pathway between the cognition and behaviour.

- When the PC or SE intervention increases the targeted cognition but an increase in behaviour is *not* observed within the individual, then the evidence does *not* provide support for the causal pathway between the cognition and behaviour.

Chapter 6 - Testing the ability of the integrated model to explain physical activity behaviour within individuals with osteoarthritis: experimental *n-of-1* studies

6.1 Abstract

Background: Evidence in support of the Theory of Planned Behaviour (TPB) is dominated by group-based studies investigating between-individual differences. The *n-of-1* (single case) design allows investigation of the TPB to explain behaviour and behaviour change *within* individuals. This study employed a series of *n-of-1* studies to investigate whether constructs from the integrated International Classification of Functioning, Disability and Health (ICF) / TPB model predicted physical activity (PA) within individuals with mobility problems due to osteoarthritis; and whether experimental investigation results in support for the candidate causal pathways within the TPB.

Method: Four experimental *n-of-1* studies were conducted. The proximal predictors of the TPB (intention, perceived controllability and self-efficacy), PA (walking and a personally identified PA) and pain were measured using diary methods for 12 weeks. For each participant, predictive and potentially causal relationships were identified in their six-week baseline data. The cognition that best predicted the PA chosen by the participant determined which theory-based intervention they received. Walking was measured by pedometer step count and personally identified PA behaviours were self-reported.

Results: The TPB predicted PA in all four participants. Action planning increased self-reported PA in one of two participants. One other participant declined the action

planning intervention yet an increase in self-reported walking was observed. The self-efficacy intervention was given to one participant; it failed to increase self-efficacy, produced a decrease in self-reported walking but did not affect objectively measured walking (pedometer step count). A weak relationship between pain and the performance of PA was identified in all individuals.

Discussion: There was support for the integrated ICF/TPB model to predict PA within individuals with osteoarthritis. Action planning received some support, whilst no support for the causal pathway between self-efficacy and PA was found. These findings highlight the need for further investigation of the predictive value and experimental application of the TPB *within* individuals. Application of the integrated ICF/TPB model will also allow the relationship between pain (impairment) and PA (activity limitation) to be further investigated.

6.2 Introduction

Osteoarthritis has been identified as the sixth most common cause of global disability (World Health Organization, 2004). In the UK, 1 in 5 of the population has osteoarthritis (Arthritis Care, 2004). Osteoarthritis causes pain and functional limitations for the individual, which impacts on their quality of life (Abell, Hootman, Zack et al., 2005). Physical activity (PA) is recognised as “a core treatment for people with osteoarthritis, irrespective of age, comorbidity, pain severity or disability” (National Institute for Health and Clinical Excellence, 2008) page 10). The European League Against Rheumatism’s evidence-based recommendations identify PA as an effective non-pharmacological treatment reducing pain and improving function in both hip and knee osteoarthritis (Jordan et al., 2003; Zhang et al., 2005). For example, a large scale clinical trial of an intervention combining calorific restriction and exercise was found to significantly improve pain, function and mobility in overweight or obese older adults with knee osteoarthritis (Messier, Loeser, Miller et al., 2004).

Regular PA is also associated with positive benefits in health-related quality of life in individuals with arthritis. Inactive men and women with arthritis were 1.2 to 2.4 times more likely to report impaired health-related quality of life than those meeting current PA recommendations (Abell et al., 2005). Furthermore, PA has also been shown to partially mediate the impact of musculoskeletal disorders, including arthritis, on quality of life in older adults (Sawatzky, Liu-Ambrose, Miller et al., 2007).

In sum, PA can play a primary and secondary preventive role in individuals with osteoarthritis. Regular PA can be used to treat the existing arthritic condition, delaying the progression of disability, and also reducing the likelihood of acquiring additional chronic illnesses associated with a sedentary lifestyle such as obesity. Older adults are those most burdened by arthritic conditions and evidence suggests that the positive health benefits associated with the modification of risk factors for disease are still obtainable later in life. Consequently, governmental health promotion strategies have highlighted the benefits of increasing PA levels in older adults (Department of Health, 2001).

Chapter 5 of this thesis provided the current evidence in support of the Theory of Planned Behaviour (TPB) and the subsequent rationale and justification for experimental investigation of the TPB within individuals. In osteoarthritis, a patient's control cognitions or beliefs about whether they can perform a behaviour have been shown to predict activity limitations (Orbell et al., 1998). Experimental studies have also shown that these control cognitions can be modified resulting in a reduction in activity limitations (Lorig et al., 2004). Further, empirical testing of Johnston's (2002) model integrating the TPB with the International Classification of Functioning, Disability and Health (ICF) has received support in the prediction of walking limitations in individuals with osteoarthritis: impairment (pain) and control cognitions independently contributed to explaining the variance in walking limitations (Dixon et al., 2008b).

Pain has been suggested to partially explain PA behaviour in individuals with osteoarthritis, specifically reducing their willingness and/or ability to engage in regular PA (McNair, Simmonds, Boocock et al., 2009). Pain experienced in osteoarthritis has been found to be intermittent, variable and transient (Allen, Coffman, Golightly et al., 2009; Gooberman-Hill et al., 2007; Focht, Ewing, Gauvin et al., 2002); thus, a study design involving frequent and repeated assessments, such as that characteristic of the *n*-

of-1 methodology, is suitable to capture the variable nature of pain perception within individuals with osteoarthritis. Within the ICF literature, pain is regarded to be an index of the theoretical construct of impairment (Cieza et al., 2004; Dreinhofer et al., 2004; Pollard et al., 2006) therefore, daily measures of pain would allow within-individual variability of the ICF construct of impairment to be identified.

Pain and PA have been described as having a bidirectional relationship in individuals with osteoarthritis: pain is a barrier to PA and a reduction in PA is associated in the long-term with weight gain, which can elevate pain perception on weight-bearing joints such as the knee and hip (Rosemann, Kuehlein, Laux et al., 2007). Furthermore, obesity is a primary risk factor for osteoarthritis (Miller, Rejeski, Williamson et al., 2003). The ICF framework similarly proposes a bidirectional relationship between the health outcomes of impairment (pain) and activity limitations (see Chapter 1 Section 1.3.1). Within the disability literature, it is more common to employ a deficit model which studies an individual's *limitation* to perform a behaviour, rather than a model which studies an individual's *actual performance* of a behaviour. However, compatible with the conceptualisation of disability as behaviour within the ICF, is the possibility to investigate disability associated with a health condition such as osteoarthritis, by measuring the performance of PA behaviour. This means that the within-individual relationship between impairment and disability in the ICF framework can be investigated employing measures of pain and PA. Therefore, the current study will test the ability of the integrated ICF/TPB model to explain PA behaviour *within* individuals with osteoarthritis.

6.2.1 Research Questions

1. Does the TPB predict PA *within* individuals with osteoarthritis?

2. Do individually tailored interventions increase PA *within* individuals with osteoarthritis?
    a) do perceived controllability and self-efficacy interventions provide support for the candidate *causal* pathway between perceived behavioural control and PA behaviour in the TPB framework?
    b) are action planning interventions effective *within* individuals with osteoarthritis?
3. What is the relationship between pain and PA *within* individuals with osteoarthritis?

6.3 Methodology

6.3.1 Design

A handheld computerised diary was used to measure the proximal predictors of the TPB, PA behaviours and pain within individuals with osteoarthritis over a period of 12 weeks. Each participant’s baseline data from 0 to 6 weeks directed the type of intervention they received: (a) a TPB-based intervention designed to increase PA by increasing either perceived controllability (PC) or self-efficacy (SE), or (b) an action planning intervention designed to increase PA, by aiding the translation of intentions into behaviour.

6.3.2 Participants

Adverts were placed in local community facilities such as the library, post office and church notice boards, asking people with knee or hip osteoarthritis to contact the researcher if they were interested in taking part in a study looking at the daily factors affecting their condition. Five people responded to the advert and were invited to take part. Four individuals accepted (80%) and gave informed consent. The sample comprised two males and two females, with an age range of 48 to 67 years. Three participants had been diagnosed with osteoarthritis of the knee(s) between 1 and 3 years earlier, the other participant had been diagnosed with osteoarthritis of the hip 2 years earlier. One female participant had a comorbid chronic pain condition of fibromyalgia.

6.3.3 Measures and Materials

An illustration of the computerised diary measures is shown in Appendix 4.

6.3.3.1 PA behaviours

Two PA behaviours were measured for each participant: ‘walking more than usual’, and a personally identified PA. Walking was assessed objectively by pedometer step count and by self-report to the question ‘How much have you walked since last filling in the diary?’ (on a sliding visual analogue scale with anchors of *less than usual* and *more than usual*, on a hidden scale from 0 to 100). Both walking measures were recorded by the participant at each diary entry. A personally identified PA behaviour, which the individual was motivated to do but limited in doing, was measured for each participant: ‘fishing’; ‘gardening’; ‘logging’; and ‘driving’ were identified. ‘Logging’ was defined as collecting, chopping and carrying logs of wood. ‘Driving’ is perhaps not typically seen to be a physically active behaviour; however, the participant explained that unpredictable pain and locking of the knee meant that she avoided driving longer distances to visit friends and family, therefore it was an important and valued disability behaviour for her. Each personally identified behaviour was assessed by self-report to the question ‘How much have you done X since you last filled in the diary?’ (on a sliding visual analogue scale with anchors of *less than usual* and *more than usual*, on a hidden scale from 0 to 100). X was replaced by the personally identified behaviour.

6.3.3.2 TPB cognitions

Intention, PC and SE cognitions in relation to each of the two PA behaviours were measured on a sliding visual analogue scale representing scores ranging from 0 to 100. The single standard TPB items with verbal anchors described in Chapter 5 Section 5.4.2 were used. Cognitions were recorded in the diary at each entry.

6.3.3.3 Other diary measures

Current pain was assessed by the question: ‘How would you describe your pain right now?’(on a sliding visual analogue scale with anchors of *no pain* and *extreme pain*, on a hidden scale from 0 to 100).

6.3.3.4 Health status measures

Health status was measured at the start of the study and after the 12-week data collection. The following measures were used:

1) SF-36 (Ware et al., 1993), described in Chapter 5 Section 5.4.3.

2) HADS (Zigmond & Snaith, 1983), described in Chapter 5 Section 5.4.3.

3) Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) (Bellamy et al., 1988), a validated self-administered questionnaire to assess symptom severity and disease-specific health-related quality of life in patients with osteoarthritis

of the knee or hip, producing normalised scores from 0 (least severe) to 100 (most severe) for pain, stiffness, physical function and a total score.

4) EQ-5D (The EuroQol Group, 1990), a standardised, non-disease specific measure of health status and health-related quality of life producing a single index score with ‘dead’ anchored at 0 and ‘full health’ at +1 (Dolan, 1997).

6.3.3.5 Materials

The diary was programmed using the software ‘Pocket Questionnaire v1.2’ (University of Aberdeen Data Management Team, 2006) and uploaded onto a handheld personal digital assistant device (Hewlett Packard iPAQ 214). Diary data was downloaded from the device to a personal computer using the Pocket Questionnaire software. The pedometer used was Omron HJ-113.

6.3.4 Procedure

Each participant was provided with a handheld diary device and instructed on its general operation including turning it on and off and charging the battery. Participants were instructed on filling in the diary and each participant completed a dummy diary entry with the researcher to ensure that they were comfortable completing it. The devices were programmed to alert the participant and prompt them to fill in the diary by ‘beeping’ them. If the participant was not able to fill in the diary immediately, the alarm could be stopped and would resound five minutes later. Participants were advised to miss the diary entry if they were not able to complete it within a 1 hour period of the original alarm. During week 6, participants were contacted to collect the devices and download the 6-week data for analyses. Devices were returned the same day before the next diary entry was due. See Chapter 5 Section 5.6 for details of the general study procedure.

6.3.5 Intervention protocol

The intervention protocol was followed as described in Chapter 5 Section 5.7. Interventions were face-to-face in participants 1, 2 and 4, but participant 3 opted for a telephone-based intervention.

6.3.6 Analyses

The general analysis protocol is detailed in Chapter 5 Section 5.8. In this study, walking behaviour was measured by self-report and pedometer step count. Therefore, when the PA selected for change by the participant was walking, intervention effects on both self-reported walking and step count are reported. In addition, this study conducted an exploratory investigation of the within-individual relationship between pain impairment and PA. This exploratory investigation was conducted on data from the baseline period because the intervention was not designed nor hypothesised to have an effect on this relationship. Consequently, after controlling for autocorrelation in the baseline data series, as described in the analysis protocol, pain series were cross-correlated with PA behaviour series. A maximum lag of one week (i.e., 14 observations) was adopted and only cross-correlations of +/-0.40 and above that also exceeded 95% confidence intervals (CI) were considered.

6.3.7 Ethics

Ethical approval was obtained from the Department of Psychology, University of Stirling.

6.4 Results

6.4.1 Description of Participants

Table 6.1 reports pre- and post-study scores from the health status assessments for each participant. SF-36 scores were evaluated with reference to age- and gender-related norms (Ware et al., 1993). HADS scores were evaluated in relation to established clinical cut-off scores whereby scores of 8 to 10 are ‘possible’ clinical disorders and scores of 11 to 21 are ‘probable’ (Zigmond & Snaith, 1983). WOMAC and EQ-5D scores were evaluated in relation to scores seen in patients awaiting joint replacement surgery, typically with moderate to severe osteoarthritis (Quintana et al., 2005, and see Table 4.2 which displays the mean scores for the 'EUROHIP' sample studied in Chapter 4).

Table 6.1 Pre- and post-study description of participants

Measure and Subscales		Participant 1		Participant 2		Participant 3		Participant 4	
		Pre	Post	Pre	Post	Pre	Post	Pre	Post
SF-36	Physical functioning	40	55	45	55	80	90	20	70
	Social functioning	62.5	62.5	75	75	87.5	100	75	87.5
	Role physical	0	0	25	0	0	75	0	0
	Role emotional	0	0	100	100	100	100	33.3	33.3
	Mental health	52	56	76	76	92	92	96	96
	Vitality	35	40	70	75	70	70	50	60
	Bodily pain	31	41	51	62	62	62	31	41
	General health	25	32	67	67	77	77	77	82
HADS	Anxiety	16	13	6	5	5	3	4	2
	Depression	15	8	4	3	3	3	8	3
WOMAC	Total WOMAC	44	46.9	17.7	17.7	27.1	31.3	69.8	41.7
	WOMAC pain	55	55	20	20	30	30	65	50
	WOMAC stiffness	62.5	62.5	37.5	37.5	37.5	50	62.5	62.5
	WOMAC function	38.2	42.6	14.7	14.7	25	29.4	72.1	36.8
EQ-5D	EQ-5D	0.62	0.62	0.8	0.76	0.69	0.8	0.69	0.8

Note. SF-36 scores range from 0 to 100 (higher score = better health state); HADS scores range from 0 to 21 (higher score = more anxious or depressed); WOMAC scores range from 0 to 100 (higher scores = more severe symptom disability); and EQ-5D scores range from -.0594 to +1 (higher score = better health state).

The descriptive findings presented in Table 6.1 show that participants had greater disability and poorer health status than age- and gender-related norms likely to be attributable to their arthritic condition; SF-36 scores for the subscales of role physical and bodily pain were lower than age- and gender-related norms, indicating worse than average health, for all participants. However, bodily pain improved in three participants from pre- to post-study assessment. Participants 1 and 4 also had lower than age- and gender-adjusted norm scores on all other SF-36 subscales. HADS anxiety and depression scores were 'normal' for three participants (participants 2, 3 & 4) but indicative of 'possible' or 'probable' clinical disorder for participant 1. However, both anxiety and depression scores reduced or remained constant over time for all four participants. WOMAC scores for participants 1, 2 and 3 were generally lower than typically seen in osteoarthritis patients awaiting joint replacement. Participant 4 had higher WOMAC scores indicating worse symptom severity but the scores did reduce considerably over time. Health related quality of life (EQ-5D) was consistently better in all participants than typically seen in patients awaiting joint replacement and participants 3 and 4 also showed an improvement from pre- to post-study assessment.

6.4.2 Diary Completion

Compliance with diary completion was high; participants completed between 91 and 100% of the possible entries. The maximum number of missed diary entries by any one participant was 16. Fourteen of these entries were during the post-intervention period, yet reasonably evenly distributed throughout, and therefore treated as missing data. There was an average of 78 diary entries per participant during the baseline period, ranging from 72 to 80 entries. In three of four participants, a short period of time elapsed between collection of baseline data and initiation of the intervention whilst the timing of the intervention was scheduled. During this period, participants continued to fill in the diary and consequently the whole pre-intervention period ranged from 72 to 86 entries with an average of 80 observations per participant. The average post-intervention period was 89 diary entries per participant, ranging from 85 to 94 entries.

6.4.3 Overall Variability in TPB Cognitions and PA Behaviours

Variability in TPB cognitions and PA behaviours over the full 12-week study period was visually inspected for each participant. For each participant, a total of 9 time plots were inspected (i.e., intention, PC and SE in relation to walking and the personally identified PA; walking step count; self-reported walking and; self-reported personally identified PA); therefore, 36 time plots were inspected in total. Variability was evident in all of them. In general, the frequency and degree of fluctuation was not uniform across cognitions or PA behaviours but rather, peculiar to the particular data series. One exception to the lack of uniformity in variability was observed in data series for PC. Seven of the 8 PC time plots (1 for each of the 2 PA behaviours studied in each of the 4 participants), generally displayed a smaller magnitude of variation (i.e. smaller fluctuations), than that observed in the time plots for all other cognitions and behaviour. Figure 6.1 illustrates this finding presenting two typical PC time plots taken from participants 3 and 4 and the two corresponding SE time plots, where substantially larger fluctuations and greater variability can be seen. The time plots display data obtained during a two-week period (i.e., 30 diary entries) and in each case a higher score represents a stronger belief.

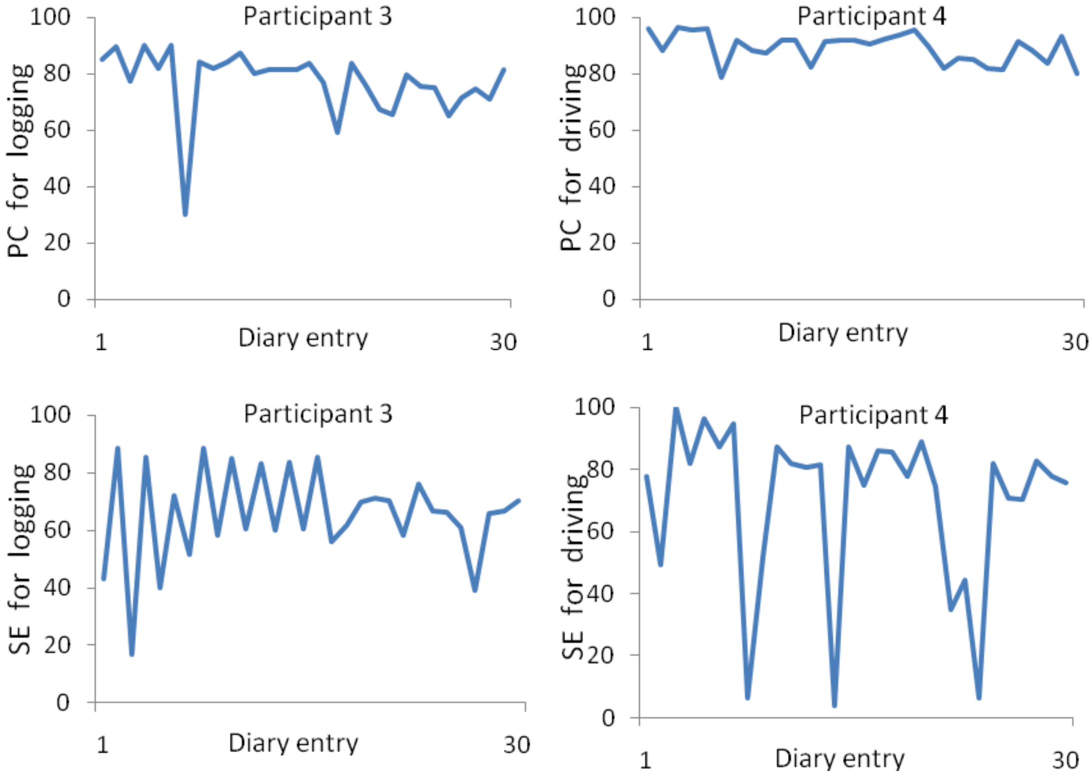


Figure 6.1 Examples of time plots of variability in PC and SE cognitions for logging and driving behaviours in participants 3 and 4 respectively

6.4.4 Baseline Data Analyses

6.4.4.1 TPB predictors of PA behaviours

At six weeks, all 36 data series displayed variability as identified by time series analyses. Twenty-eight series (78%) revealed significantly correlated scores across successive days and were therefore pre-whitened. First-order autoregressive relationships (i.e., lag of 1 diary entry) were the most common, being found in 13 series, followed by second-order relationships (i.e., lag of 2 entries) in 17 series. The remaining 8 series displayed autocorrelation at lags ranging from 3 to 9. Table 6.2 details the cross-correlations and time lags between each TPB cognition and the PA selected for change by each participant, using pre-whitened variables as appropriate.

Table 6.2 Baseline cross-correlations and time lags between TPB cognitions and the physical activity selected for intervention, and type of intervention given

Chosen Behaviour	Intention	PC	SE	Intervention
Participant 1				
Walking				
Self-report	-.35 lag 2[b]	*ns*	-.37 lag 0[b]	SE
Steps	.44 lag -7	*ns*	-.436 lag 0[a]	
Participant 2				
Walking				
Self-report	.50 lag 1	*ns*	.28 lag 0[b]	Action Plan
Steps	.52 lag 1[a]	.31 lag -12[b]	*ns*	
Participant 3				
Walking				
Self-report	.44 lag 1	-.28 lag -7[b]	*ns*	Action Plan[c]
Steps	.57 lag 1[a]	*ns*	*ns*	
Participant 4				
Driving				
Self-report	.61 lag 1[a]	*ns*	.25 lag 1[b]	Action Plan

[a]The highest cross-correlation of those ≥.40(+/-); [b]Cross-correlations <.40 (+/-) but exceeding 95% CI; [c]Participant 3 chose not to form an action plan.

6.4.4.2 Intervention selection

TPB cognitions predicted the PA selected for change for all four participants; therefore, none of the participants received an intervention by default but rather each participant was offered an individually tailored intervention based on the cognition that best predicted the PA they chose. The type of intervention given to each participant is shown in the final column of Table 6.2. Participant 1 chose walking as the behaviour they wanted to do more of, and SE was the strongest predictor of their walking therefore participant 1 received the SE intervention in relation to walking. The other three participants were given the action planning intervention; for these participants intention was the strongest predictor of their chosen PA behaviours.

However, in one case (participant 3) the offered action planning intervention was declined. This participant received the verbal explanation that based on the information he provided in the diary during the baseline period, his intention predicted his walking behaviour. Yet, when he was asked to create an action plan, he declined saying that after spending his working life living by a routine, he was no longer inclined to make plans as a retired man. Thus, even though an action plan was not formed by this participant, he continued to fill in the diary for the remaining weeks of the study and therefore, his data was analysed in accordance with the analysis protocol in the same way as the other participants' data. Analyses of participant 3's data are presented and discussed separately to the results for participants 2 and 4, who accepted the action planning intervention and completed the task of creating an action plan.

6.4.5 Effect of SE Intervention for Participant 1

6.4.5.1 Effect on SE and the chosen PA behaviour

Time plots of the SE measure and chosen PA across the 12-week study period were visually inspected for intervention effects (see Appendix 5). Tests comparing pre- and post-intervention scores and the corresponding effect sizes are reported in Table 6.3. The SE intervention for this participant did not change SE or walking measured by pedometer step count; however, a significant decrease in self-reported walking was observed. The post-intervention relationship between SE and walking is also shown in Table 6.3. Baseline relationships are also displayed to facilitate comparison. The post-intervention association between SE and walking behaviour (self-report and steps) is very similar to that identified before the intervention; there is a moderately strong negative relationship between SE and walking with a time lag of 0 (i.e., lower SE predicted more walking reported since last diary entry: $r = -.47$ self-report and $r = -.39$ steps, lag 0).

Table 6.3 Effect of self-efficacy (SE) intervention for walking for participant 1

Intervention Effect	
Effect on SE score	
Mean pre-intervention (*StE*)	39 (2.0)
Mean post-intervention (*StE*)	37 (1.3)
Statistic *t*	.99
Effect size *d*	*ns*
Effect on self-report walking	
Mean pre-intervention (*StE*)	48 (2.3)
Mean post-intervention (*StE*)	40 (1.7)
Statistic *t*	2.9**[b]
Effect size *d*	.47
Effect on number of steps	
Median pre-intervention (*n*)	840 (79)
Median post-intervention (*n*)	1404 (69)
Statistic *U*	2714.5
Effect size *d*	*ns*
Post-intervention SE – behaviour relationship	
Self-report[a]	-.47 lag 0 **(-.37 lag 0)**
Steps[a]	-.39 lag 0 **(.44 lag 0)**

Note. Cognitions and self-reported PA were measured on visual analogue scales from 0-100 (a higher score = a stronger cognition and more PA respectively).
StE = standard error; *$p < .05$. **$p < .01$. ***$p < .001$.
[a]Figures in **bold** and parentheses denote the baseline SE–behaviour relationships; [b]Significant *decrease* in mean self-reported walking

6.4.6 Effect of Action Planning Intervention for Participants 2 and 4

6.4.6.1 Effect of action planning on the chosen PA behaviour

Time plots of the chosen PA in each participant across the 12-week study period were visually inspected (see Appendix 5). Tests of the intervention effect on PA and the corresponding effect sizes for the chosen PA of walking for participant 2 and for the chosen PA of driving for participant 4 are reported in Tables 6.4 and 6.5 respectively. The action planning intervention increased the chosen PA in one individual (participant 4), as measured by self-report.

Table 6.4 Effect of action planning intervention for walking for participant 2

Intervention Effect	
Effect on self-report walking	
Mean pre-intervention (*StE*)	46 (2.3)
Mean post-intervention (*StE*)	48 (2.0)
Statistic *t*	-.66
Effect size *d*	*ns*
Effect on number of steps	
Median pre-intervention (*n*)	1840 (70)
Median post-intervention (*n*)	2623 (83)
Statistic *U*	2686.5
Effect size *d*	*ns*
Post-intervention intention – behaviour relationship	
Self-report[a]	.35 lag 1 **(.50 lag 1)**
Steps[a]	.47 lag 1 **(.52 lag 1)**

Note. Self-reported PA was measured on visual analogue scales from 0-100 (a higher score = more PA); *StE* = standard error.
[a]Figures in **bold** and parentheses denote the baseline intention–behaviour relationships

Table 6.5 Effect of action planning intervention for driving for participant 4

Intervention Effect	
Effect on self-report driving	
Median pre-intervention (*n*)	6 (79)
Median post-intervention (*n*)	18 (87)
Statistic *U*	2397.0**
Effect size *d*	.54
Post-intervention intention – behaviour relationship	
Self-report[a]	.39 lag 1 **(.61 lag 1)**

Note. Self-reported PA was measured on visual analogue scales from 0-100 (a higher score = more PA); $*p < .05$. $**p < .01$. $***p < .001$.
[a]Figures in **bold** and parentheses denote the baseline intention–behaviour relationships

6.4.6.2 Effect of action planning on the intention-behaviour relationship

Post-intervention relationships between intention and walking for participant 2 and between intention and driving for participant 4 are also displayed in Tables 6.4 and 6.5, respectively. Baseline relationships are also displayed to facilitate comparison. Baseline and post-intervention associations between intention and the chosen PA (as measured by self-report and step count for walking, and self-report for driving) were positive and with a lag of 1 for both participants (i.e., stronger intention predicted more PA 12 hours later). However, the strength of the post-intervention relationships is weaker than the baseline relationships for both participants; reducing from .50 and .52 to .35 and .47 for self-reported walking and steps, respectively for participant 2, and reducing from .61 to .39 for self-reported driving for participant 4.

6.4.7 Analyses Relating to Participant 3 Who Declined the Action Planning Intervention

Table 6.6 Pre- and post-intervention data for participant 3 who declined the action planning intervention on walking

Uncompleted Intervention Effect	
Effect on self-report walking	
Mean pre-intervention (*StE*)	32 (1.8)
Mean post-intervention (*StE*)	44 (2.1)
Statistic *t*	4.11***
Effect size *d*	.64
Effect on number of steps	
Median pre-intervention (*n*)	3145 (81)
Median post-intervention (*n*)	2177 (85)
Statistic *U*	3379.5
Effect size *d*	*ns*
Post-intervention intention – behaviour relationship	
Self-report[a]	*ns* **(.44 lag 1)**
Steps[a]	.37 lag 3 **(.57 lag 1)**

Note. Self-reported PA was measured on visual analogue scales from 0-100 (a higher score = more PA); *StE* = standard error; $^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$.

[a] Figures in **bold** and parentheses denote the baseline intention–behaviour relationships

Even though participant 3 declined the action planning intervention and did not create an action plan, an increase was observed in his chosen PA of walking, as measured by self-report; no increase was observed in the objective walking measure of pedometer step count. With respect to the relationship between intention and behaviour, the significant association between intention and self-reported walking evident at baseline was not present post-intervention. However, the association between intention and steps at baseline continued to be present post-intervention but the lag increased from a lag of 1 at baseline to a lag of 3 post-intervention (i.e., stronger intention predicted more steps 36 hours later: $r = .37$, lag 3). Furthermore, similar to the decrease in the strength of association between intention and behaviour identified in participants 2 and 4, a decrease in the strength of the association from .61 at baseline to .39 at post-intervention was identified in participant 3.

6.4.8 Relationship Between Pain and PA During the Baseline Period

Table 6.7 Descriptive statistics of pain scores, cross-correlations and time lags between pain and physical activity (PA) behaviours at baseline

PA Behaviours	Mean Pain (StE)	Correlation and Time Lag
Participant 1	34 (1.3)	
Walking (self-report)		-.38 lag -9
Steps		-.25 lag -3
Fishing		*ns*
Participant 2	31 (1.9)	
Walking (self-report)		*ns*
Steps		*ns*
Gardening		.29 lag -5
Participant 3	28 (1.0)	
Walking (self-report)		*ns*
Steps		*ns*
Logging		.25 lag 3
Participant 4	60 (2.0)	
Walking (self-report)		.28 lag -6
Steps		.29 lag 0
Driving		*ns*

Note. Pain was measured on a visual analogue scale from 0-100 (a higher score = more pain); *StE* = standard error; All cross-correlations <.40 (+/-) but exceeded 95% CI

The pain data series for each participant displayed variability as identified by time series analyses. The pain series for three of four participants revealed significantly correlated scores across successive days and were therefore pre-whitened. First-order autoregressive relationships (i.e., lag of 1 diary entry) were evident in two series and a second-order relationship in one series. Table 6.7 details cross-correlations and time lags between pain and each of the PA behaviours for each participant, using pre-whitened variables as appropriate. None of the relationships between pain and any of the PA behaviours for any of the four participants met the criterion for significance, i.e., cross-correlations of +/-0.40 and above that also exceeded 95% CI.

6.5 Discussion

In accordance with the criteria to evaluate the TPB defined in Chapter 5, Section 5.8.1, it can be concluded that this series of *n-of-1* studies found some support for the application of the TPB pathways in PA within individuals with osteoarthritis; intention predicted PA in all four participants and either PC or SE, as the component parts of perceived behavioural control, predicted at least one of the PA behaviours measured in all four participants. The SE intervention for participant 1, failed to increase SE, produced a decrease in self-reported walking but did not affect objectively measured walking by pedometer step count. Action planning increased self-reported PA in one of two participants, but no objective measure of PA was available. The action planning intervention was declined in the remaining participant, yet an increase in self-reported but not objectively measured walking was observed.

6.5.1 Is There Support for the TPB to Predict PA Within Individuals with Osteoarthritis?

Within-individual variability in the proximal predictors of behaviour as proposed by the TPB (Ajzen, 1991) and all PA behaviours was observed in all participants over the 12-week period. The presence of variability allowed within-person investigation to see whether an individual was more likely to engage in PA at times when they had stronger intentions and perceived behavioural control than at other times. Intention predicted PA during the baseline period in all participants. Intention was positively correlated with the chosen behaviour with a time lag of one diary entry in three of four individuals; a stronger intention predicted a longer duration of PA 12 hours later. This identified direction and close temporal nature of the relationship between intention and PA is consistent with the contention that intention is a *proximal* predictor of behaviour, as proposed by the TPB. In the remaining individual (i.e., participant 1), intention was significantly predictive of objectively measured walking (i.e., step count) but not self-reported walking; intention and step count were positively related with a lag of -7, which indicates that a stronger intention predicted more steps 84 hours *earlier*. The positive sign of this relationship between intention and PA behaviour is consistent with the TPB; a stronger intention predicts more PA. However, the identified temporal lag suggests that *earlier* walking predicted *later* intention. The precedence direction of this relationship conflicts with the possible causal pathway between intention and behaviour assumed by the TPB and therefore, this finding cannot be easily explained by the TPB. Replication of this finding is needed to test its soundness.

With respect to whether an individual was more likely to engage in PA at times when they had stronger perceived behavioural control than at other times, PC did not significantly predict PA in any individual. The related construct of SE, however, did predict PA in one individual (i.e. participant 1); a negative relationship between SE and step count was identified with a lag of 0, indicating that at times when the individual reported having walked more steps, he also reported feeling less confident about walking within the next 12 hours. In addition, the same negative relationship with a lag of 0 was identified between SE and the self-reported measure of walking, albeit that the cross-correlation of -.39 fell short of the +/-.40 criterion for significance.

The TPB predicts that stronger SE will result in more PA, whereas the current finding suggests that more PA (walking) results in weaker SE. Further, this negative relationship between SE and PA is counter to self-efficacy theory which would predict that a successful mastery experience performing a behaviour would increase, not decrease, SE to perform the behaviour (Bandura, 1977). Studies within the learning literature have similarly identified a negative relationship between SE and task or exam performance at the within-individual level of analysis, and have suggested that personal goals and goal level (difficulty) may help to explain the finding (Vancouver, Thompson, & Williams, 2001; Vancouver & Kendall, 2006). It is possible that the participant in the current study had conflicting goals. For example, an individual with osteoarthritis, for whom walking is difficult and pain is common, may on occasion possess the goal to control pain rather than to be active. They may feel that after having walked more than normal, they are not confident of their ability to do a lot more walking because their current goal is to control pain by **not** being active. A measure of SE over control of pain in future *n-of-1* studies of individuals with mobility problems would permit further exploration of this possible explanation.

An *n-of-1* study of a patient with the disabling condition of chronic idiopathic axonal polyneuropathy also identified a negative relationship between SE and walking on the same day (Schroder, 2008). However, unlike the current study which found that more *performance of* walking was associated with lower SE, Schroder's (2008) study found that more self-reported *limitations in* walking was associated with lower SE over walking (cross correlation of -.49 lag 0) and that performance of walking measured by pedometer step count was not associated with SE. Additional within-individual assessment of the intricate nature of the relationship between SE and walking is needed. Measuring walking both in terms of the *performance of* and *limitations in* will help to develop an evidence base from which conclusions can be made about the relationship between SE and walking. The majority of applications of the TPB require *a priori* assumptions about the causal lag between constructs or the direction of causal flow (Sutton, 2003). However, applying an *n-of-1* methodology and within-individual time series analyses is suitable to explore the temporal precedence of SE and other TPB constructs, and behaviour, without such assumptions.

In sum, intention was a strong predictor of PA in all participants, whilst the evidence for the role of perceived behavioural control was more mixed. SE was a strong predictor of PA in one participant, whilst PC did not strongly predict PA in any participant. The finding that SE is a better predictor than PC has been shown in earlier group-based studies of PA (Dzewaltowski, Noble, & Shaw, 1990; Terry & O'Leary, 1995). In addition, a predictive relationship between SE and PA, and one between PC and PA was only seen in one participant; only one or other of SE and PC predicted PA in the other three participants. This finding lends further support to the evidence and argument for PC and SE as separable constructs (Ajzen, 2002; Trafimow, Sheeran, Conner et al., 2002).

The second key issue addressed by this study was whether individually tailored interventions increase PA *within* individuals with osteoarthritis, lending support to the candidate causal pathway between perceived behavioural control and behaviour in the TPB or lending support to action planning interventions.

6.5.2 Is there Support for a Causal Pathway Between SE and PA Behaviour Within Individuals with Osteoarthritis?

Unfortunately, the SE intervention did not increase SE in participant 1, therefore one cannot draw conclusions about the potential causal relationship between SE and PA. Previous evidence has shown that SE and control beliefs can be experimentally manipulated and concurrent reductions in disability can be seen in populations similar to that investigated in the current study. For example, SE has been increased in arthritic populations resulting in reductions in disability and improvements in function (Lorig et al., 2004) and control beliefs have been increased in a chronic pain sample, using the experimental manipulation on which the current intervention was based, resulting in reductions in disability (Fisher & Johnston, 1996). Thus, it is unclear why the current intervention failed to increase SE in participant 1; study replication is needed in other individuals to be able to draw conclusions about the effectiveness of this intervention.

6.5.3 Is there Support for Action Planning Within Individuals with Osteoarthritis?

There is support for action planning as a behaviour change technique to increase PA *within* one of two participants who completed the intervention. A medium-sized positive effect was identified in the self-reported PA behaviour of driving for participant 4. No objective measure of this behaviour was available. In contrast, the action planning intervention was not supported within participant 2, for whom no change in either self-reported walking or steps was identified. The temporal nature of the relationship between intention and PA for both participant 2 and 4 was unchanged from baseline to post-intervention; stronger intention predicted more PA 12 hours later, lending further support to intention as a proximal predictor of behaviour *within* individuals.

Participant 3 declined the action planning intervention, i.e., he did not make an action plan. Previous studies employing action planning interventions have found that within the intervention group, participants who actually make the action plan are more likely to perform the target behaviour than participants who do not make the plan (Michie et al., 2004; Rutter et al., 2006). Thus, on the basis previous studies, an increase in PA in participant 3 may not have been expected. However, a medium-to-large effect, based on Cohen's (Cohen, 1992) criteria of effect sizes, in self-reported walking was identified, but no concurrent increase in objectively measured step count was found.

A possible explanation for the identified increase in self-reported walking, even though the participant did not make an action plan, is that he received feedback on his data. In accordance with the intervention protocol, the participant was told that one of the things that influenced his walking behaviour was his intention and that making an action plan would help him to carry out his intentions. It is possible that the feedback component of the intervention may have generated the increase in self-reported walking in this participant. Michie et al's (2008) classification of behaviour change techniques has concluded that the technique of providing feedback can be effective in changing the construct domains of 'beliefs about capabilities' and 'beliefs about consequences' and 'motivation and goals', which are believed to overlap with domains of 'self-efficacy' and 'anticipated outcomes/attitude' and 'intention' identified by Fishbein et al (2001, cited by Michie, Johnston, Abraham et al., 2005). Inspection of intention and SE for walking in this participant revealed that neither intention nor SE increased from pre- to post-intervention, suggesting that the increase in self-reported walking cannot be attributed to a change in either TPB cognition (see Appendix 5 for additional analyses of the *n-of-1* data for this study). Neither beliefs about consequences or attitude were measured precluding further examination of this potential mechanism.

An alternative explanation for the significant increase in self-reported walking in participant 3 may relate to demand characteristics. Participants were told that they would receive individualised feedback to help them increase performance of a chosen PA. Therefore, it is possible that, even though participant 3 did not make the action plan, he may have wanted to please the researcher and appear to have increased his walking, which may have led him to self-report more walking behaviour in the post-intervention period. However, if demand characteristics were an issue in this case then a significant increase in step count may also have been expected which was not the case; although pedometer step count is an objective measure of walking, participants were required to record the number of steps displayed on the pedometer in the diary at each entry and therefore if the participant had wanted to appear to have walked more, he could have falsified the number of steps he recorded to reflect this. As has been previously acknowledged in the PA literature, recall of walking is poor and self-reported and pedometer step counts do not always coincide (Scott, Eves, French et al., 2007). In the current study, providing feedback relating to walking and offering the action planning intervention to participant 3 may have heightened his awareness and attention to walking behaviour, which in turn produced an increase in the subjective measure of walking without actually increasing the number of steps taken.

If the effectiveness of action planning interventions relies on the participant actually making a plan, then knowing why an individual may be unwilling to do so is very important. Participant 3 said that he did not want to create an action plan because he had spent his working life living by a prescribed routine. This reason may be common to many retired individuals and therefore future applications of action planning interventions in retired samples should establish the acceptability of action planning in this population.

6.5.4 What is the Relationship Between Pain and PA Within Individuals with Osteoarthritis?

The pain series for three of four participants displayed a significant autocorrelation in pain intensity across successive days. This finding concurs with that previously identified in the daily study of pain within individuals with arthritis (Affleck, Tennen, Urrows et al., 1991). In the current study, none of the cross-correlations between pain and any of the PA behaviours for any of the four participants were +/-0.40 or above and that also exceeded 95% CI. This finding is similar to that of Dixon et al (2008b) who found that pain explained less variance in walking limitations than did the TPB cognitions of intention and perceived behavioural control in a group study of individuals with osteoarthritis.

With respect to the relationships between pain and PA which fell below the +/-0.40 criterion yet did exceed 95% CI, time series analyses revealed that the relationship was different for each individual and was also not consistent across PA behaviours within the same individual. For example, there was variation in the direction and size of the lags between pain and PA, and pain did not predict both walking and the personally-identified PA in any of the participants. Focht et al (2002) studied the relationship between PA and pain in people with knee osteoarthritis using an ecological momentary assessment technique, where participants made between 5 and 6 diary entries per day, and found that acute exercise resulted in a significant increase in pain; however the observed increase in pain after PA was transient and had reduced later that day. If increases in pain are as short-lived as reported in Focht et al’s (2002) study then it is possible that the time period of 12 hours between diary entries adopted in the current study, may have masked more temporary increases in pain and that more frequent diary entries are needed to capture more transient changes in pain.

To summarise, pain was weak predictor of PA within individuals however the exact nature of the relationship remains unclear and further within-individual investigation is warranted. Unlike the majority of disability research, the current study did not employ a deficit model examining an individual’s limitation to perform a behaviour, but rather focused on actual performance of PA. In a large community-based study, pain and general health impairment directly predicted activity limitations in individuals reporting chronic pain, providing support for the pathway between impairment and activity limitations as proposed by the ICF framework (Johnston, Dixon, Hannaford et al., 2009). However, unlike the current study’s finding that pain weakly predicted PA within individuals with osteoarthritis, Johnston et al’s (2009) study reported that pain and general health impairment did not directly predict actual walking behaviour in their sample. Further research is needed to fully understand the intricate nature of the relationship between pain and disability, conceptualised both in terms of a deficit model assessing what an individual does not do (activity limitations) and in terms of what an individual does do (performance of PA).

There are strengths and limitations with this study. A strength of this study was the use of computerised diaries which prevent the participant from back-filling earlier missed entries and/or reviewing earlier entries, improving the reliability of the results and reducing reporting bias (Stone & Shiffman, 2002). Diary compliance was very good suggesting that the diary methodology was well received by participants. Participant feedback confirmed this, with all participants reporting that the diary was easy and quick to fill in and that this was aided by having the opportunity to negotiate timings to make entries at the beginning of the study. Further anecdotal observations from this study are discussed in Chapter 8 Section 8.5.

A limitation of this study was that participants were recruited from the community for their interest in learning about the daily factors that affect their condition, thus, the sample was self-selecting and cannot claim to be representative of the sample population. However, the *n-of-1* design examines within-individual differences and therefore, this study makes no claim to generalise findings beyond the individual that was studied. A second limitation was that three of four participants chose walking as the PA selected for change allowing an objective measure of pedometer step count to be used in analyses of the effect of the intervention. However, pedometer accuracy is consistently found to be reduced in slower walking paces (Tudor-Locke, Williams, Reis et al., 2002) and the number of steps may be underestimated in elderly populations with gait disorders (de Bruin, Hartmann, Uebelhart et al., 2008). This issue raises doubt over the reliability of steps measured in the current study as participants had lower limb osteoarthritis. Nonetheless, the within-individual methodology used in the current study investigated variability in walking over time *within* the same individual and therefore, any factor relating to the accuracy of step count would have been present throughout the full study period. Thus, it is less of an issue in the current study.

A third limitation similarly relates to the use of pedometers. As previously discussed, participants were asked to record the number of steps in the diary at each entry which means that the reliability of the number of steps recorded is limited by the potential for participants to falsify the number they recorded. This limitation could be overcome by employing more sophisticated objective measures of walking and PA such as accelerometers, which do not display the amount of PA the participant has performed and can store many days of data, which the researcher can download at the end of the study period.

Finally, the failure of the SE intervention to increase SE in the participant who received this intervention, may indicate that the intervention could have been better designed. Ongoing work to classify behaviour change techniques according to theoretical constructs and to determine the most effective techniques (Abraham & Michie, 2008; Michie et al., 2008; Michie, Abraham, Whittington et al., 2009) could be used to identify candidate behaviour change techniques to be tested in future *n-of-1* studies.

6.5.5 Summary

In sum, there was evidence in support of the integrated ICF/TPB model to predict PA *within* individuals with osteoarthritis; the TPB predicted PA in all four participants and there was some evidence in support of a weak relationship between pain and PA. This study also provides some evidence in support of action planning interventions to increase self-reported PA *within* some individuals with osteoarthritis. However, future work is needed to replicate and consolidate findings with objective measures of PA. Further research investigating whether making the action plan is intrinsic to the effectiveness of the planning intervention is also needed. No evidence in support of the possible causal pathway between SE and PA was found. The implications of these findings are discussed more fully in conjunction with the findings of the study reported in Chapter 7 (see section 7.5.6).

Chapter 7 - Testing the ability of the Theory of Planned Behaviour to explain physical activity within healthy individuals: experimental *n-of-1* studies

7.1 Abstract

Background: Evidence in support of the utility of the Theory of Planned Behaviour (TPB) in the prediction of physical activity (PA) is dominated by group-based designs investigating between-individual differences. Whether or not the TPB can be used to explain PA behaviour and behaviour change *within* individuals is yet to be established. The *n-of-1* (single case) design allows behaviour change within individuals to be studied. This study employed a series of *n-of-1* studies to test the ability of the TPB to explain PA behaviour and PA behaviour change within healthy individuals.

Method: Six experimental *n-of-1* studies were conducted. The TPB cognitions (intention, perceived controllability and self-efficacy) and PA (walking, gym and a personally identified PA) were measured using diary methods for 12 weeks. In each participant, predictive and potentially causal relationships were identified in their six-week baseline data. Each participant specified the PA they wanted to increase, and the cognition that best predicted that PA determined which theory-based intervention they received. When no cognition was identified a default intervention was given. Walking was measured by pedometer step count, gym behaviour was self-reported and validated by electronic membership records, and personally identified PA behaviours were self-reported.

Results: The TPB predicted PA in three of six participants. The individually tailored perceived controllability intervention increased perceived controllability and increased

self-reported and objectively measured PA in one of two participants. The self-efficacy intervention given to one participant by default did not increase self-efficacy but an increase in objectively measured walking was observed. Action planning did not increase PA in any of the three participants who received the action planning intervention.

Discussion: There was some support for the ability of the TPB to predict PA within healthy individuals. However, support for the candidate causal pathways within the TPB was weaker. Insufficient patient motivation may account for the null finding in relation to action planning within individuals. Further investigation of the predictive and explanatory utility of the TPB *within* individuals is needed. Moreover, future work investigating PA within healthy individuals should consider employing the 'healthy' version of the integrated ICF/TPB model.

7.2 Introduction

Investigation into physical activity (PA) behaviours in healthy individuals has tended to focus on the cognitive determinants of PA performance such as those proposed by the Theory of Planned Behaviour (TPB). For example, Hagger et al's (2002) review of the TPB in PA research concluded that TPB cognitions can account for approximately 29% of the variance in PA behaviour. Chapter 5 of this thesis provided the background and rationale for the experimental investigation of the TPB within individuals and Chapter 6 reported the findings of a series of *n-of-1* studies within individuals with osteoarthritis. As a comparative experimental study, a series of *n-of-1* studies in healthy individuals will allow investigation of whether the same theory and methodology applies within a different population.

There is significant empirical evidence for the positive health benefits associated with regular PA (Department of Health, 2004b; Salmon, 2000; World Health Organization, 2003). However, the literature on the effectiveness of interventions to increase PA has shown that even though there is some evidence in support of interventions to produce short-term changes in PA, there is less evidence in support of the maintenance of these changes (Hillsdon, Foster, Cavill et al., 2005). Thus, experimental investigation of the TPB in PA *within* individuals will provide a more comprehensive understanding of individual behaviour change, which can be used to develop theory-based interventions that may be more successful in changing and maintaining PA.

7.2.1 Research Questions

1. Does the TPB predict PA *within* healthy individuals?

2. Do individually tailored interventions increase PA *within* healthy individuals?

c) do perceived controllability and self-efficacy interventions provide support for the candidate *causal* pathway between perceived behavioural control and PA behaviour in the TPB framework?

d) are action planning interventions effective *within* healthy individuals?

7.3 Methodology

7.3.1 Design

A web-based or paper diary was used to assess the proximal predictors of the TPB and PA behaviours within healthy individuals over a period of 12 weeks. Each participant's baseline data from 0 to 6 weeks directed the type of intervention they received: (a) a TPB-based intervention designed to increase PA by increasing either perceived controllability (PC) or self-efficacy (SE), or (b) an action planning intervention designed to increase PA, by aiding the translation of intentions into behaviour.

7.3.2 Participants

New recruits to the University of Stirling fitness centre were approached by the researcher after an induction session with centre staff. A total of seventeen people were invited to take part in the study and seven accepted (41.2%), gave informed consent and were screened with the Physical Activity Readiness Questionnaire (PAR-Q; (Canadian Society for Exercise Physiology, 2002). The original sample comprised three males and four females, with an age range of 24 to 71 years. One male participant aged 25 years regularly failed to complete the diary and dropped out of the study at week 5, providing insufficient data to support analyses. Therefore, the final sample was six individuals, two males and four females. Five were members of the public, one was a postgraduate student and one was a member of university staff. Participants were remunerated with £50 on completion of the study.

7.3.3 Measures and Materials

An example of the diary measures is shown in Appendix 6.

7.3.3.1 PA behaviours

Three PA behaviours were measured for each individual: ‘walking more than usual’, ‘working out in the gym’ and a personally identified PA. Walking was assessed objectively by pedometer step count (Omron HJ-113) and recorded by the participant at each diary entry. Gym behaviour was assessed by self-report to the question ‘Have you worked out in the gym since you last filled in the diary?’ (*yes* or *no*). When an affirmative response was given, the participant was asked to report the duration (in minutes) and intensity (on a 10-point verbal rating scale from 1 *no exertion* to 10 *maximal exertion*). Gym behaviour was objectively validated by entry to the gym via an electronic membership card system. A personally identified non-gym based PA behaviour was measured for each individual: ‘walking the dog’; ‘going for a walk’; ‘doing the exercise DVD’; and for two participants, ‘swimming’. Each of these behaviours was assessed by self-report in the same way as gym behaviour i.e., an affirmative answer to having performed the behaviour was followed by two questions relating to the duration and intensity of performance. When the personally identified behaviour was ‘walking the dog’ and ‘going for a walk’, self-reported behaviour was supplemented by pedometer step count as an objective measure of general walking behaviour. Like gym behaviour, swimming attendance was also objectively validated by the electronic card system. One participant could not specify a personally identified PA and recorded walking and gym behaviour only.

7.3.3.2 TPB cognitions

Intention, PC and SE cognitions in relation to each of the three PA behaviours were measured on a 10-point verbal rating scale using the single standard TPB items with verbal anchors as described in Chapter 5 Section 5.4.2. Cognitions were recorded in the diary at each entry.

7.3.3.3 Health status, fitness and physiology measures

Health status, fitness and physiology measures were taken at the start of the study and at the end of the 12-week data collection period. Fitness and physiology measures were also taken at study weeks 5 and 9. The following measures were taken:

1) SF-36 (Ware et al., 1993), described in Chapter 5 Section 5.4.3.

2) HADS (Zigmond & Snaith, 1983), described in Chapter 5 Section 5.4.3.

3) Leisure Time Exercise Questionnaire (LTEQ) (Godin & Shephard, 1985), a self-administered questionnaire to assess the amount of exercise performed during leisure time, producing a metabolic equivalent (MET) score.

4) Chester step test (Stevens & Sykes, 1996), a test of aerobic fitness predicting maximal aerobic power (VO_2max).

5) EUROFIT assessment of health-related fitness (Council of Europe Committee for the Development of Sport and UKK Institute for Health Promotion Research, 1995)

including average grip strength (kg), broad jump (cm), sit-ups (number in 30 sec), push-ups (number in 30 sec) and sit and reach flexibility (cm).

6) Blood pressure (mmHg) and Anthropometric measures of body mass index (BMI), percentage body fat and waist-to-hip ratio.

7.3.3.4 Materials

Participants could choose to use web-based or paper diaries depending on personal preference and internet accessibility. The web-based diary was held on a website linked to the University of Stirling main website (http://www.alittlelab.stir.ac.uk) and data was collected on a server maintained in the Department of Psychology. Participants filled in the diary using an anonymised login known only to the researcher. Date and time of filling in the diary was automatically recorded. The paper diary was an A5 sized booklet. A space was provided at the beginning of each entry for the participant to record the date and time. The pedometer used was Omron HJ-113.

7.3.4 Procedure

SMS message reminders were sent twice daily prompting diary completion. Participants filled in the diary online or in paper format. In addition to the pre- and post-study assessments, participants also met with the researcher at weeks 5 and 9 for additional fitness assessments. Chapter 5 Section 5.6 details the general study procedure.

7.3.5 Intervention Protocol

The intervention protocol was followed as described in Chapter 5 Section 5.7. Interventions were delivered by telephone in five participants; one participant opted to receive the intervention face-to-face. Each participant was required to choose one behaviour that they were motivated to do more frequently or for longer periods, however, one participant was unable to identify just one such behaviour choosing both gym and swimming; consequently, this individual received two interventions, one for each PA.

7.3.6 Analyses

The analysis protocol is detailed in Chapter 5 Section 5.8. In this study, generic walking behaviour was measured by pedometer step count. Therefore, when the PA selected for change by the participant was a personally identified PA involving walking behaviour, intervention effects on this behaviour measured by self-reported duration and pedometer step count were investigated.

7.3.7 Ethics

Ethical approval was obtained from the Department of Psychology, University of Stirling.

7.4 Results

7.4.1 Description of Participants

Table 7.1 reports pre- and post-study scores from the health status and fitness assessments for each participant. SF-36 scores were evaluated with reference to age- and gender-related norms (Ware et al., 1993). HADS scores were evaluated in relation to established clinical cut-off scores whereby scores of 8 to 10 are ‘possible’ clinical disorders and scores of 11 to 21 are ‘probable’ (Zigmond & Snaith, 1983). Aerobic capacity (VO_2 max) was converted into age- and gender-related norm fitness ratings of *excellent*, *good*, *average*, *below average* or *poor* (Stevens & Sykes, 1996). An increase in aerobic capacity of 4 mlsO_2/kg/min or more is said to reflect a significant improvement in VO_2max (Buckley, Sim, Eston et al., 2004). The LTEQ metabolic equivalent (MET) scores, the EUROFIT measures and the anthropometric measures provide an indication of the general fitness of the participants. Normal blood pressure is less than 140/90 but not less than 90 systolic and 60 diastolic. A ‘normal’ BMI ranges from 18.5 to 24.9; ‘overweight’ ranges from 25 to 29.9 and ‘obese’ is 30 or more.

Table 7.1 Pre- and post-study description of participants

Measure and Subscales		Participant 1		Participant 2		Participant 3		Participant 4		Participant 5[a]		Participant 6	
		Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
SF-36	Physical functioning	85	100	100	100	95	90	90	90	100	-	85	85
	Social functioning	60	70	100	100	88	25	100	100	100	-	100	88
	Role physical	100	100	100	100	100	75	100	100	100	-	100	75
	Role emotional	67	100	33	100	100	0	100	100	33	-	100	100
	Mental health	50	84	44	72	56	20	84	84	60	-	60	55
	Vitality	45	67	55	70	65	25	45	40	40	-	60	55
	Bodily pain	100	74	61	84	84	84	51	84	72	-	41	84
	General health	87	97	95	87	66	92	67	52	72	-	37	47
HADS	Anxiety	8	8	12	5	6	0	11	14	6	-	6	4
	Depression	5	0	4	1	0	0	3	3	4	-	5	4
LTEQ	Leisure exercise MET	16	56	41	57	35	25	33	23	36	-	0	20
Chester step	VO_2 max mlsO_2/kg/min	43	47	51	56	39	39	34	37	44	43	39	39
EUROFIT	Grip strength kg	27	28	35	36	89	95	23	19	35	39	27	19
	Broad jump cm	89	95	111	116	124	132	95	90	82	102	94	93
	Sit-ups in 30 sec	11	15	18	26	16	23	11	17	25	20	9	15
	Push-ups in 30 sec	11	15	7	20	16	13	10	14	17	26	10	12
	Sit and reach cm	4	3	-11	-12	-1	1	4	3	-20	-16	1	6
Anthropometric	Blood pressure mmHg	130/85	107/73	123/69	153/77	124/65	108/61	115/77	114/73	151/81	136/76	102/66	121/82
	BMI	27	26	26	27	21	21	32	33	23	23	38	37
	Body fat %	37	34	21	22	22	22	60	59	4	5	68	60
	Waist-to-hip ratio	0.85	0.81	0.92	0.89	0.68	0.64	0.78	0.82	0.93	0.94	0.78	0.74

Note. SF-36 scores range from 0 to 100 (higher score = better health state); HADS scores range from 0 to 21 (higher score = more anxious or depressed); higher LTEQ MET score, higher VO_2max, and higher EUROFIT scores indicate more leisure time exercise, greater aerobic capacity and better fitness, respectively; higher BMI, body fat % and waist-to-hip ratio indicate more obese. [a]Participant 5 post-study assessment was not conducted therefore SF-36, HADS and LTEQ scores are missing; post-study fitness measures reported were taken during week 9.

To summarise the findings presented in Table 7.1, the health status and fitness assessments indicate that participants were generally healthy without any significant impairment. Aerobic capacity consistently corresponded to an *excellent* fitness rating in three participants (participant 1, 2 & 5). Further, a significant increase in mlsO_2/kg/min was identified in two of these participants (participant 1 & 2). Aerobic capacity consistently corresponded to an *average* fitness rating in the other three participants (participant 3, 4 & 6). Four participants (participant 1, 2, 4 & 6) generally had SF-36 scores that matched or exceeded age- and gender-adjusted norms, whilst some SF-36 subscales fell within the 25th percentile for participants 3 and 5. HADS depression score was ‘normal’ for all participants, but HADS anxiety score was indicative of ‘possible’ clinical disorder for participant 1 and ‘probable’ clinical disorder for participants 2 and 4; however, anxiety did reduce to ‘normal’ in participant 2 from pre- to post-study assessment. All participants had normal range blood pressure. Two participants had a ‘normal’ BMI (participants 3 & 5), two participants were ‘overweight’ (participants 1 & 2) and two participants were ‘obese’ (participants 4 & 6). Four participants (participants 1, 2, 5 & 6) showed improvements in some of the EUROFIT and anthropometric measures over the course of the 12-week study period.

7.4.2 Diary Completion

Compliance with diary completion was high with participants completing between 96 and 100% of the possible entries. The maximum number of missed diary entries by any one participant was 7 and these were distributed across the 12-week period and therefore treated as missing data. Self-reported gym and swimming behaviours coincided with electronic membership card records 92% of the time. No record was available to validate the self-reported behaviour for the other 8% of the time due to entry system malfunction.

There was an average of 79 diary entries per participant during the baseline period, ranging from 71 to 97 entries. The study protocol only required a 6-week baseline period; however, this was extended in the individual for whom 97 entries were recorded due to a delay in the retrieval of their paper diary. In all participants, a short period of time elapsed between collection of baseline data and initiation of the intervention whilst the timing of the intervention was scheduled. During this period, participants continued to fill in the diary and consequently, the whole pre-intervention period ranged from 105 to 131 entries with an average of 117 observations per participant. The average post-intervention period was 72 diary entries per participant, ranging from 48 to 92 entries. The participant for whom only 48 entries were recorded reflects early termination of the study due to illness. It has been recommended that a minimum of 50 data observations exist within each data series to establish a reliable time series model of the data (Morley, 1994). This suggested minimum was exceeded in each of the baseline and post-intervention series for all individuals with the exception of the series containing 48 observations. This shorter series was nonetheless analysed as per protocol; the associated limitations in these analyses are addressed in the discussion.

7.4.3 Overall Variability in TPB Cognitions and PA Behaviours

Variability in TPB cognitions and PA behaviours over the full 12-week study period was visually inspected for each participant. For each of the five participants who reported on three PA behaviours, a total of 14 time plots were inspected (i.e., 9 cognition plots for intention, PC and SE in relation to walking, gym and the personally identified PA; and 5 PA plots for walking step count, gym duration, gym intensity, personally identified PA duration and personally identified PA intensity). Only nine plots were available for the participant with no personally identified PA. Consequently, 79 plots were inspected in total. Variability in TPB cognitions and PA behaviours was evident for all participants over the 12-weeks. Nevertheless, the frequency and degree of fluctuation was not uniform across cognitions or behaviours, but rather was unique to that data series. Figure 7.1 illustrates typical types of variability observed in the cognition measures. The time plots display data obtained over a two-week period (i.e., 30 diary entries) and in each case a higher score represents a stronger belief (e.g., higher intention).

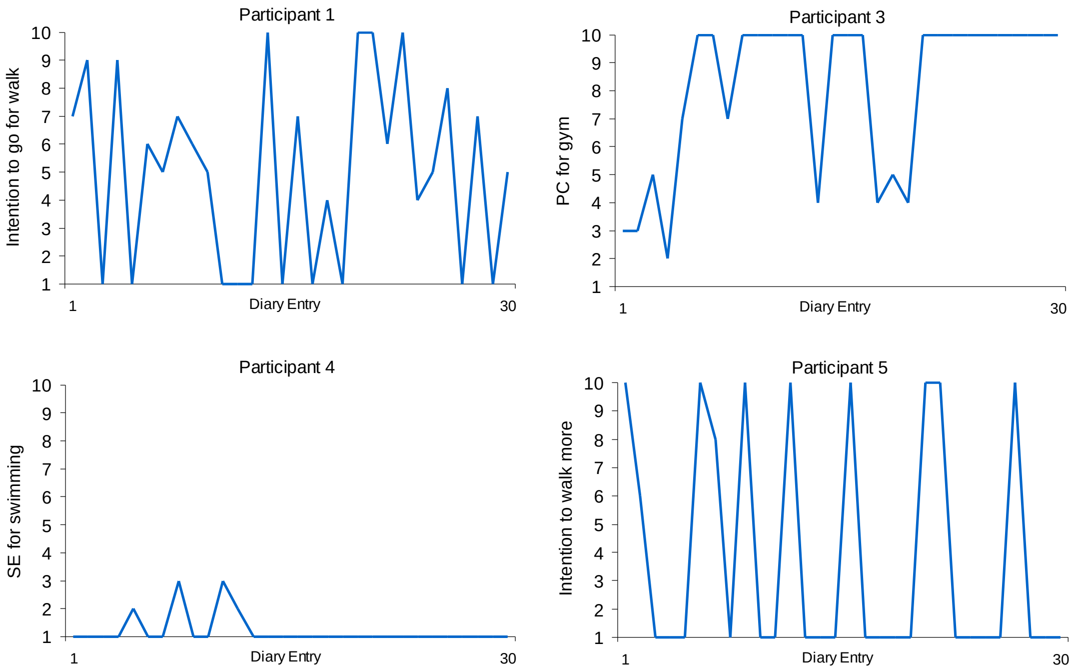


Figure 7.1 Examples of time plots of variability in TPB cognitions for PA behaviours illustrating various types of variability: frequent fluctuations (participant 1), ceiling effects (participant 3), floor effects (participant 4) and extreme response set (participant 5).

7.4.4 Baseline Data Analyses

7.4.4.1 <u>TPB predictors of PA behaviours</u>

At six weeks, 75 of the 79 data series displayed variability as identified by time series analyses. Sixty-one series (81%) revealed significantly correlated scores across successive days and were therefore pre-whitened. First-order autoregressive relationships (i.e., lag of 1 diary entry) were the most common, being found in 17 data series, followed by second-order relationships (i.e., lag of 2 entries) in 14 series. The remaining 30 data series displayed autocorrelation at lags ranging from 3 to 14. Table 7.2 details the cross-correlations and time lags between each TPB cognition and the PA selected for change by each participant, using pre-whitened variables as appropriate. Measures of self-reported duration and intensity of gym and personally identified PA were consistently highly correlated in all participants (*r* > .90 in all); consequently, only duration of these behaviours is reported.

Table 7.2 Baseline cross-correlations and time lags between TPB cognitions and the physical activity selected for intervention, and type of intervention given

Chosen Behaviour	Intention	PC	SE	Intervention
Participant 1				
Go for a walk				
Minutes	.50 lag 1	.53 lag 1[a]; .34 lag 0[b]	.50 lag 1	PC
Steps	*ns*	*ns*	*ns*	
Participant 2				
Walk the dog				
Minutes	*ns*	*ns*	*ns*	Action Plan
Steps	*ns*	*ns*	*ns*	
Participant 3				
Gym				
Minutes	.36 lag 4[b]	*ns*	.27 lag -1[b]	Action Plan
Swimming				
Minutes	.36 lag 1[b]	*ns*	.28 lag -2[b]	Action Plan
Participant 4				
Gym				
Minutes	.63 lag 1	.64[a] lag 1	.62 lag 1	PC
Participant 5				
Walking				
Steps	*ns*	*ns*	*ns*	SE
Participant 6				
Gym				
Minutes	.54 lag 1[a]; .39 lag 3[b]	.40 lag 1; .46 lag 3	.46 lag 1; .41 lag 3	Action Plan

Note. Participants 1 and 2 selected their personally identified PA for change and these behaviours involved walking, therefore both self-reported minutes and step count measures of behaviour are available. Participant 5 chose general walking behaviour which was not self-reported therefore only step count is available.

[a]The highest cross-correlation of those ≥.40(+/-). [b]Cross-correlations <.40 (+/-) but exceeding 95% CI.

7.4.4.2 Intervention selection

The type of intervention given to each participant is shown in the final column of Table 7.2. TPB cognitions predicted the PA selected for change for three of six participants (participants 1, 4 & 6), therefore each of these participants was offered an *individually tailored intervention* based on the cognition that best predicted the PA they chose. PC was the strongest predictor of the chosen PA behaviours for participants 1 and 4; these participants received the PC intervention. Intention was the strongest predictor of the chosen PA for participant 6; this participant received the action planning intervention. None of the TPB cognitions predicted the chosen PA and fulfilled the significance criterion for participants 2, 3 and 5, therefore each of these participants was offered the *default intervention* of action planning as specified in the intervention protocol. Participants 2 and 3 accepted the action planning intervention; however, after discussion with participant 5, it was decided that a SE intervention would be more suitable[2].

7.4.5 Effect of PC or SE Intervention for Participants 1, 4 and 5

7.4.5.1 <u>Effect of the individually tailored PC intervention on PC and the chosen PA for participants 1 and 4</u>

Time plots of the PC measure and chosen PA across the 12-week study period for each participant were visually inspected for intervention effects (see Appendix 7). Tests comparing pre- and post-intervention scores and the corresponding effect sizes are reported in Table 7.3. An abrupt increase in PC and a somewhat more gradual increase in PA were observed after the PC intervention for participant 1 (see Figure 7.2). This pattern was supported by a significant increase in PC and the behaviour ‘going for a

---

[2] Participant 5 chose walking behavior when asked which PA he was most motivated to do more frequently or for longer periods, however he clarified this choice and said he had no desire to increase walking but rather was motivated to continue walking as much as he currently did. The action planning strategy is said to function at the post-intentional action stage increasing performance of behavior via the enactment of intentions (Gollwitzer, 1993); however, as the individual was not motivated to increase current performance of behavior, asking him to form an action plan designed to increase behavior seemed redundant. Hence, an intervention to increase control beliefs was judged to be a more suitable default intervention to allow him to maintain his current walking level. Visual inspection of the time plot for PC over walking revealed measurement ceiling effects (i.e., feeling in ‘complete control’ over 90% of the time). Inspection of the SE time plot however, revealed variability typical of extreme response set with SE fluctuating from ‘not at all confident’ to ‘extremely confident’. The ceiling effects identified in PC would make the detection of an increase in PC difficult and therefore it was decided that the individual would receive the SE rather than PC intervention.

walk' measured by self-reported duration and objectively measured pedometer step count. In contrast, the PC intervention for participant 4 led to a significant decrease in PC and no change in behaviour.

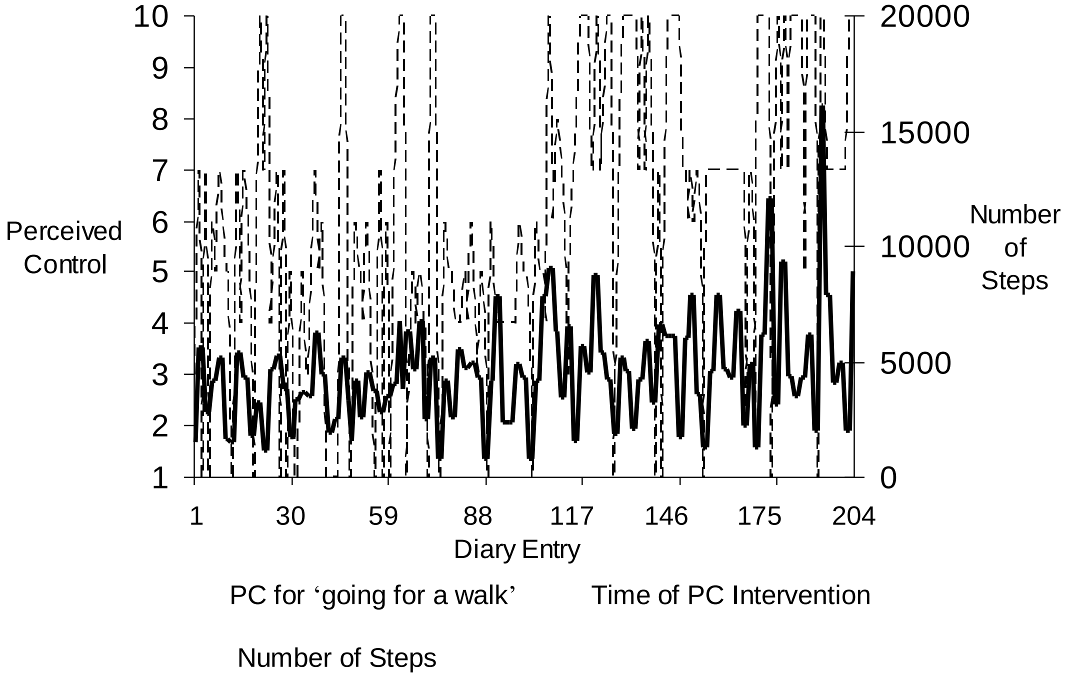


Figure 7.2 Effect of perceived controllability (PC) intervention on the chosen behaviour of 'going for a walk' (measured by pedometer step count) for participant 1.

Table 7.3 also displays post-intervention relationships between PC and the chosen PA in each participant. Baseline relationships are also displayed to facilitate comparison. The association between higher PC and longer duration of PA at pre-intervention for participant 1 is somewhat weaker post-intervention and only seen with a time lag of 1 (i.e., higher PC predicted longer duration of PA 12 hours later: *r* = .27, lag 1). Minimal variability in PC post-intervention for participant 4 prohibited computation of cross-correlations and the possibility to observe any deviation from the positive association identified pre-intervention between gym behaviour and PC 12 hours earlier *(r* = .64 lag 1).

Table 7.3 Effect of perceived controllability (PC) intervention for participants 1 and 4

Intervention Effect	Participant 1 ‘go for walk’	Participant 4 ‘gym’
Effect on PC score		
Median pre-intervention (*n*)	5 (110)	1 (131)
Median post-intervention (*n*)	7 (92)	1 (71)
Statistic *U*	1948.0***	3337.0***[b]
Effect size *d*	.61	.75
Effect on minutes of chosen PA		
Median pre-intervention (*n*)	0 (110)	0 (131)
Median post-intervention (*n*)	20 (91)	0 (71)
Statistic *U*	3483.5***	4399.0
Effect size *d*	1.28	*ns*
Effect on number of steps		
Mean pre-intervention (*StE*)	3986 (179.5)	
Mean post-intervention (*StE*)	5135 (296.9)	
Statistic *t*	3.31***	
Effect size *d*	.55	
Post-intervention PC – behaviour relationship		
Minutes[a]	.27 lag 1 **(.53 lag 1; .34 lag 0)**	**- (.64 lag 1)**
Steps[a]	*ns* **(*ns*)**	

Note. Cognitions were measured on 10-point verbal rating scale (a higher score = a stronger cognition). Dashes indicate cross-correlations could not be computed; *StE* = standard error; **p* < .05. ***p* < .01. ****p* < .001.

[a] Figures in **bold** and parentheses denote the baseline PC–behaviour relationships; [b] Significant *decrease* in median PC score

7.4.5.2 Effect of the default SE intervention on SE and the chosen PA for participant 5

Time plots of the SE measure and chosen PA across the 12-week study period for participant 5 were visually inspected for intervention effects (see Appendix 7). Tests comparing pre- and post-intervention scores and the corresponding effect sizes are reported in Table 7.4. The SE intervention did not modify SE but a highly significant increase in walking measured by pedometer step count was observed. The post-intervention relationship between SE and walking is also displayed in Table 7.4. Baseline relationships are also displayed to facilitate comparison. No significant cross-correlations between SE and walking were identified pre-intervention, however, a strong negative association was identified post-intervention (i.e., lower SE predicts more steps reported since last diary entry: $r = -.70$, lag 0).

Table 7.4 Effect of self-efficacy (SE) intervention for walking for participant 5

Intervention Effect	Participant 5
Effect on SE score	
Median pre-intervention (*n*)	1 (131)
Median post-intervention (*n*)	1 (48)
Statistic *U*	3057.0
Effect size *d*	*ns*
Effect on number of steps	
Mean pre-intervention (*StE*)	4798 (218.0)
Mean post-intervention (*StE*)	7531 (632.0)
Statistic *t*	4.09***
Effect size *d*	1.07
Post-intervention PC – behaviour relationship	
Steps[a]	-.70 lag 0 (***ns***)

Note. Cognitions were measured on 10-point verbal rating scale (a higher score = a stronger cognition). *StE* = standard error; *p < .05. **p < .01. ***p < .001.
[a] Figures in **bold** and parentheses denote the baseline SE–behaviour relationships

7.4.6 Effect of Action Planning Intervention for Participants 2, 3 and 6

7.4.6.1 Effect of the individually tailored action planning intervention on the chosen PA and the intention-behaviour relationship for participant 6

No effects of action planning on the chosen PA across the 12-week study period for participant 6 were observed from visual inspection of the time plots (see Appendix 7). This finding was supported by tests comparing pre- and post-intervention gym scores, which indicated no significant difference between pre- and post- intervention gym scores (Table 7.5).

Table 7.5 Effect of action planning intervention for gym behaviour for participant 6

Intervention Effect	Participant 6
Effect on minutes of gym behaviour	
Median pre-intervention (*n*)	0 (116)
Median post-intervention (*n*)	0 (79)
Statistic *U*	4460.0
Effect size *d*	*ns*
Post-intervention intention – behaviour relationship	
Minutes[a]	.41 lag 1 **(.54 lag 1; .39 lag 3)**

Note. [a]Figures in **bold** and parentheses denote the baseline intention–behaviour relationship

Post-intervention relationships between intention and gym behaviour for participant 6 are also displayed in Table 7.5. Baseline relationships are also displayed to facilitate comparison. The association between intention and longer duration of gym behaviour at baseline for participant 6 was weaker post-intervention and only seen with a time lag of 1 (i.e., the lag 1 relationship was $r = .54$ at baseline and $r = .41$ post-intervention, and the lag 3 relationship at baseline was not identified at post-intervention).

7.4.6.2 Effect of the default action planning intervention on the chosen PA and the intention-behaviour relationship for participants 2 and 3

Visual inspection of the time plots of the chosen PA across the 12-week study period for participants 2 and 3 did not identify any effect of the action planning intervention (see Appendix 7). Tests comparing pre- and post-intervention PA scores confirmed this finding, failing to detect any significant difference between pre- and post- intervention PA scores (Table 7.6).

Table 7.6 Effect of action planning intervention for participants 2 and 3

Intervention Effect	Participant 2 ‘walk dog’	Participant 3 ‘gym’	 ‘swim’
Effect on minutes of PA			
Median pre-intervention (*n*)	0 (105)	0 (109)	0 (109)
Median post-intervention (*n*)	0 (71)	0 (78)	0 (78)
Statistic *U*	3717.0	4140.5	4072.0
Effect size *d*	*ns*	*ns*	*ns*
Effect on number of steps			
Mean pre-intervention (*StE*)	4668 (279.7)		
Mean post-intervention (*StE*)	5069 (264.1)		
Statistic *t*	.99		
Effect size *d*	*ns*		
Post-intervention intention – behaviour relationship			
Minutes[a]	.48 lag 1 **(*ns*)**	*ns* **(.36 lag 4)**	*ns* **(.36 lag 1)**
Steps[a]	.29 lag 1 **(*ns*)**		

Note. Participant 2 selected their personally identified PA that involved walking; therefore, results relating to the behaviour measured in self-reported minutes and step count are reported. Participant 3 chose two PA behaviours, gym and swimming. *StE* = standard error.
[a]Figures in **bold** and parentheses denote the baseline intention–behaviour relationships

Post-intervention relationships between intention and the chosen PA for participant 2 and 3 are also reported in Table 7.6. Baseline relationships are also displayed to facilitate comparison. Post-intervention, associations were detected for participant 2 that were not evident at baseline; both self-reported duration of walking the dog and step count were associated with intention (i.e., stronger intention predicted longer duration of walking and more steps 12 hours later: $r = .48$ and $r = .29$ respectively, lag 1). For participant 3, no significant associations between intention and either PA behaviour (gym or swimming) were identified post-intervention, in contrast to the associations identified at baseline (i.e., at baseline stronger intention predicted longer duration of gym behaviour 48 hours later: $r = .36$, lag 4, and; longer duration of swimming 12 hours later: $r = .36$, lag 1).

7.5 Discussion

In accordance with the criteria to evaluate the TPB defined in Chapter 5, Section 5.8.1, it can be concluded that this series of *n-of-1* studies provide support for the application of the TPB pathways within three of six healthy individuals; the TPB cognitions predicted PA in three participants; these individuals received individually tailored interventions. Further, the individually tailored PC intervention increased PC and PA in one participant providing support in favour of a causal pathway between PC and PA. However this effect was not seen in the other participant receiving the individually tailored PC intervention failing to provide support for the causal pathway. The individually tailored action planning intervention for participant 6 did not increase PA. There was no support for the TPB in the remaining three individuals; none of the TPB cognitions predicted PA and these individuals received default interventions. The default SE intervention given to one participant did not increase SE; however, a highly significant increase in objectively measured behaviour was observed. In accordance with the pre-defined criteria this finding does not provide support for a causal pathway between SE and PA. The default action planning intervention did not increase PA in either of the two remaining participants.

7.5.1 Is There Support for the TPB to Predict PA Within Healthy Individuals?

During the baseline period the TPB predicted PA within three of six individuals (i.e., participants 1, 4 & 6). Intention, PC and SE were all positively correlated with the self-reported amount of the chosen PA with a time lag of one diary entry in these participants; a higher cognition score predicted a longer duration of PA 12 hours later. These findings provide support for the TPB to predict PA within *some* individuals. In addition, the close temporal nature of the identified relationships, with a time lag of approximately 12 hours, is consistent with the TPB cognitions being *proximal* predictors of behaviour.

However, there was no support for the predictive utility of the TPB within the remaining three participants (i.e., participants 2, 3 & 5). None of the measured TPB cognitions predicted the chosen PA in any of these participants. In the case of participant 5, this finding may reflect the extreme response set pattern observed in intention and SE, and ceiling effects noted in PC. It is possible that the item measurement scales were not suitable to capture the variability in cognitions for this individual, thus decreasing the likelihood of detecting significant associations with PA. However, this explanation is less plausible for participants 2 and 3, for whom the TPB also failed to predict PA. In these participants, scores on all cognitions spanned the full range of the item scales suggesting that, rather than a lack of measurement sensitivity, the TPB was not applicable in these individuals.

The second key issue addressed by this study was whether individually tailored interventions increase PA *within* healthy individuals, lending support to the candidate causal pathway between perceived behavioural control and behaviour in the TPB or lending support to action planning interventions. Participants 1, 4 and 6 received individually tailored interventions

7.5.2 Individually Tailored Interventions

7.5.2.1 Is there support for a causal pathway between PC and PA behaviour within healthy individuals (participants 1 and 4)?

Findings from participant 1 provide support for a causal pathway between PC and PA behaviour within an individual. The PC intervention produced positive medium-to-large effects (based on Cohen's 1992 criteria of effect sizes) on the individual’s control beliefs and walking behaviour, measured by self-reported minutes and objectively measured pedometer step count. The magnitude of this effect is not dissimilar to Fife-Shaw et al’s (2007) statistical simulations, which showed that a ‘medium’ size (0.5 *SD*) increment in perceived behavioural control engendered a significant increase in the proportion of the sample who acted.

The identified relationship between PC and walking behaviour for participant 1 showed that stronger PC resulted in longer duration of PA 12 hours *later* providing further support for a possible causal pathway between perceived behavioural control and PA, as suggested by the TPB. However, even though an increase in objectively measured pedometer step count was identified following the PC intervention, there was no evidence of a predictive relationship between PC for ‘going for a walk’ and pedometer step count either before or after the intervention. This limits interpretation of the findings as direct support of a causal pathway between PC and *objectively measured* walking. Pedometer step count measured ‘general’ walking behaviour whereas self-reported minutes of walking measured the specific behaviour of ‘going for a walk’; therefore, the lack of a predictive relationship between PC for ‘going for a walk’ and step count may simply reflect the fact the behaviour in question is subtly different. Future *n-of-1* studies should ensure that there is no discrepancy between the subject of the cognition and the behaviour.

In contrast to participant 1, the PC intervention for participant 4, designed to increase PC for gym behaviour, resulted in a significant reduction in PC and no concurrent change in behaviour. During the baseline period there was evidence in support of the TPB within this individual, whereby PC was identified as a strong predictor of gym behaviour 12 hours later; therefore, it is difficult to infer that the TPB was not applicable in this individual and for that reason the intervention did not produce the anticipated effects. The intervention was individually tailored and based on a previously successful experimental manipulation of control beliefs (Fisher & Johnston, 1996); however, it is possible that in spite of the theoretical basis for the intervention, the specific behaviour change techniques employed were not suitable or effective for this individual. An alternative and plausible explanation for the ineffectiveness of the PC intervention for this participant may relate to the fact that this individual changed gyms one week after the intervention and yet continued to fill in the diary with respect to the original gym, which perhaps explains the observed decrease in PC. Thus, the findings from this participant do not provide any support for a causal pathway between PC and behaviour.

7.5.2.2 Is there support for action planning within healthy individuals (participant 6)?

Contrary to previous empirical support for action planning as a technique that can have a medium-to-large positive effect on the performance of many behaviours (Gollwitzer & Sheeran, 2006), the individually tailored action planning intervention in the current study did not produce a significant increase in the chosen PA behaviour in participant 6. This finding can be interpreted in several ways. First, it is possible that action planning is not an effective technique to increase PA *within* individuals. In the same way as investigation of the TPB has been dominated by group-based studies, investigation of the effectiveness of action planning has similarly been dominated by group-based studies investigating change in behaviour in a group receiving an action planning intervention compared with a control or other intervention type group (Luszczynska, 2006; Sheeran & Silverman, 2003). As previously discussed, such between-person analyses fall prey to the problem of averaging effects of the group obscuring identification of the effect on any one individual. It may be that action planning is effective for some individuals but not with others, but that the averaging effect in group designs masks this variability; participant 6 in the current study may be an example of an individual for whom action planning is not effective.

An alternative explanation for the ineffectiveness of the action planning intervention for participant 6 may lie in the relationship between goal intentions and action planning within this individual. Several studies have found that the relationship between the formation of an action plan and behaviour is moderated rather than mediated by intention strength, such that an individual’s intention needs to be high in order for action planning to successfully engender behaviour change (Norman & Conner, 2005; Sheeran, Webb, & Gollwitzer, 2005). Pre-intervention intention scores for participant 6 fell at the lower end of the 10-point rating scales with a median score of 1, indicating levels of motivation that may have been insufficient for an action planning intervention. When strong positive intentions do not exist, a motivational intervention to increase goal intentions may be more suitable (Sheeran et al., 2005). The intervention protocol required participants to identify a PA behaviour they **wanted** to increase. Therefore, it was assumed that participant 6 was motivated and possessed the goal intention necessary for action planning. However, this assumption was incorrect and future *n-of-1* studies should ensure that participant’s goal intention is strong enough to support action planning.

Interactions between goal intentions and action planning can be studied in group-based deigns to see whether participants for whom the intervention was successful were also those with stronger intentions (Sheeran et al., 2005). However, the within-individual design and the small sample size in the current study preclude full moderation analyses of the relationship between intention and behaviour. Similarly, other authors have shown that an individual's level of perceived behavioural control in conjunction with their intention may affect the outcome of action planning interventions (Rutter et al., 2006). Again this could not be tested in the current study. In order to test hypotheses such as these, multilevel modelling analyses could be conducted across a series of *n-of-1* trials testing the effectiveness of action planning within many individuals.

A final observation regarding the effect of the action planning intervention for participant 6 was that significant increases in all TPB cognitions were identified (all *p*s<.01; see Appendix 7 for additional analyses of the *n-of-1* data for this study). This finding is at odds with the action planning literature that acknowledges that action planning does not affect behaviour by enhancing cognitions, in the sense that creating a plan does not affect an individual's motivation to perform the behaviour (Milne et al., 2002; Sheeran et al., 2005). In Chapter 6 Section 6.5.3, post hoc analyses did not find support for the possibility that the feedback component of the action planning intervention increased one of the TPB cognitions, which mediated the increase in self-reported walking identified in participant 3 of that study. However, in the current study, the identified increases in all TPB cognitions in participant 6 suggest that in this individual the feedback component may have increased one or more of the construct domains, which contain the TPB constructs of intention PC and SE and that have been identified as being influenced by feedback (Michie et al., 2008).

In summary, three participants received individually tailored interventions based on their baseline data; there was support for the PC intervention in one of two participants and no support for action planning in the remaining participant.

7.5.3 Default Interventions

7.5.3.1 Is there support for a causal pathway between SE and PA behaviour within healthy individuals (participant 5)?

There was no direct evidence in support of a causal relationship between SE and PA because despite failing to increase SE in participant 5, there was an increase in walking. Indeed, it is possible that the SE measure was insensitive to any changes in SE produced by the intervention as previously discussed. Analyses of the post-intervention data identified a strong negative relationship between SE and PA such that at times when the individual reported having walked more steps, he also reported feeling less confident about walking within the next 12 hours. The same negative relationship was identified between SE and objectively measured and self-reported walking in one individual with osteoarthritis in the study reported in Chapter 6 (see section 6.5.1). As previously discussed, this negative relationship is counter to self-efficacy theory which would predict that a successful mastery experience performing a behaviour would increase not decrease SE to perform the behaviour (Bandura, 1977). However, it was suggested in Section 6.5.1 that conflicting goals may help to explain the negative relationship between SE and walking in individuals with mobility problems, who may on occasion choose to control their pain rather than be active. In the current study, participants were recruited on the basis of being healthy and a new recruit to the fitness centre; however, during conversations with the researcher, participant 5 reported having arthritic symptoms of pain and stiffness in his knees but that as of yet he had not sought a medical diagnosis. Therefore, it is possible that the identified negative relationship between SE and walking for this participant may similarly reflect the possibility that after having walked more than normal, the participant's confidence of their ability to do a lot more walking is lower because their current goal is to control pain by **not** being active. It is also noted that this finding may be spurious due to the fact that this post-intervention data series had marginally less data observations than is recommended for time series analyses (i.e., this series contained 48 observations and 50 observations is the recommended minimum number of observations per data series to establish a reliable time series model of the data (Morley, 1994).

Replication of the finding of a negative relationship between SE and walking in an arguably 'healthy' individual is needed to test the robustness of the finding. Moreover, further examination of the exact nature of the relationship between SE and walking in individuals with mobility problems including a measure of control cognitions for managing pain would be useful.

7.5.3.2 Is there support for action planning within healthy individuals (participants 2 and 3)?

There was no support for action planning in either of the participants who received this intervention by default; no significant increase in the chosen PA was identified in participant 2 or participant 3. As discussed in the interpretation of the failure of the action planning intervention to increase PA in participant 6 (see section 7.5.2.2), the current evidence for action planning in the literature is dominated by findings from group-based designs which report findings that may not reflect the effect of any one individual. Therefore, similarly to participant 6, participants 2 and 3 may also be exemplar individuals for whom action planning is not an effective behaviour change technique. Furthermore, similarly to participant 6, pre-intervention intention scores for participants 2 and 3 were at the lower end of the 10-point rating scales with median scores of 3 and 1 respectively. Therefore, it is possible that these participants may not have possessed the goal intention required for action planning.

It should be noted that for participant 3, the lack of success of the action planning intervention to increase either gym or swimming behaviour may be explained by the occurrence of a personal event. Two weeks after the intervention, participant 3 was advised against doing any PA by her doctor because of health problems. Inspection of scores on the SF-36 health-related quality of life measure lends support to this explanation: six of eight SF-36 domains decreased from pre- to post-study (see Table 7.1). This explanation is further supported by the identified significant decreases in intention for each type of PA from pre- to post intervention ($p<.05$: see Appendix 7).

In summary, action planning was not successful in changing PA in any of the three participants who received this intervention in the current study; neither the participant who received the individually tailored action planning intervention nor the participants who received the action planning intervention by default increased their chosen PA. However, there is some support for the candidate causal pathway between perceived behavioural control and PA behaviour within individuals. Experimental manipulation of PC increased PC and PA in one of two participants receiving individually tailored PC interventions, and even though the default experimental manipulation of SE did not increase SE, an increase in step count was identified. It is argued that a lack of measurement sensitivity in relation to the SE measure limited the identification of an increase in SE.

There are some limitations of this study. First, it is possible that the self-reported cognitions measures were not sensitive enough to detect variability in cognitions which may explain the identified floor and ceiling effects, and extreme response set patterns. This could have been overcome by further individualisation of the study by developing measurement items tailored to each participant, which may have improved the ability to detect response variability within each individual. Unfortunately, all participants filled in the same online diary and therefore, items were standardised across participants. Secondly, the current recommendations are to take 10,000 steps a day (British Heart Foundation, 2009), however, for participant 1 the PC intervention only produced an increase in the mean number of steps only from 3986 steps at pre-intervention to 5135 steps at post-intervention. This suggests that despite the statistically significant increase in steps, future work is needed to improve the PC intervention in order to increase PA enough to meet the current recommended levels.

7.5.4 Summary

In sum, there was evidence in support of the TPB to predict PA behaviour *within* some healthy individuals. There was also evidence in support of the casual pathway between PC and PA *within* a healthy individual, whilst the evidence for a causal pathway between SE and PA *within* a healthy individual was weak but may reflect a lack of measurement sensitivity. Finally, there was no evidence in support of action planning *within* healthy individuals; however, this null finding may have been the result of insufficient patient motivation. Study replication with attention to the sensitivity of measurement items and participant’s motivational status would be useful.

This study tested the ability of the TPB to explain PA within healthy individuals. Chapter 6 tested the ability of the integrated model of disability, which integrates the

TPB with the International Classification of Functioning, Disability and Health (ICF), to explain PA within individuals with osteoarthritis. The ICF is a model of health outcomes with a healthy functioning arm and a disability arm (see Chapter 1 Section 1.3.1). ‘Body function & structure’ is the healthy counterpart to ‘impairment’, which features in the disability arm. Therefore, future *n-of-1* studies could test the ‘healthy’ version of the integrated ICF/TPB model in healthy individuals. This would require a measure of ‘body function & structure’ available to daily objective assessment and/or self-report.

7.5.5 Summary and Implications of the Findings of the Studies Reported in Chapters 6 and 7

The TPB can predict PA behaviour *within* individuals with osteoarthritis and *within* some healthy individuals. The TPB-based interventions were not universally successful in changing either the targeted cognition or increasing PA. Potential explanations and interpretations of the specific findings of each study are discussed in the respective chapters however some general conclusions and implications follow.

First, the *n-of-1* design is suitable to test theory *within* individuals as recognised in the MRC Complex Interventions Guidance (Craig et al., 2008) and can be used to test theoretical models such as the TPB. Second, individually tailored interventions can be employed to target potentially causal determinants of behaviour and advance knowledge of the specific causal pathways in the TPB. Third, behaviour change interventions that target predictive cognitions can be tested *within* individuals. It is important to acknowledge that these series of *n-of-1* studies adopted pre-defined criteria to evaluate theory. However, employing different criteria would mean that the findings from these studies would lead to different conclusions regarding whether the evidence supports or does not support the theory under investigation. For example, these studies specified that when one or more of the theoretical constructs predicted behaviour in an individual, then the evidence was said to provide support for the TPB. This means that support for the TPB was found even if not all of the TPB constructs were identified as predicting behaviour. Arguably, a more rigorous evaluation of the TPB would insist that *all* TPB constructs must predict behaviour for there to be support for the theory. Ongoing work and discussion about such issues in theory development within the behavioural sciences is needed (Abraham & Michie, 2008; Michie & Prestwich, 2010).

In contrast to group-based studies which assess between-individual differences, the focus of an *n-of-1* study is to examine within-individual variation. As a consequence, findings from different individuals are not intended to be compared nor are findings from one individual intended to be generalisable to other individuals from the same population. That said, the lack of consistent findings between the participants in terms of the predictive ability of the TPB; the effect of individually-tailored interventions on TPB constructs and PA behaviour; and the effect of individually tailored interventions versus default interventions, suggest the presence of individual differences.

Studies employing group-based designs which pool data from all individuals in the group are likely to mask these individual differences. In contrast, within-individual analyses identify person-specific variability and can detect individual response to an intervention (Molenaar & Campbell, 2009). More within-individual investigation will mean that subgroups within a population will become evident. This will improve the design of interventions tailored to the characteristics of subgroups within the target population. For example, future work could conduct multiple *n-of-1* studies in order to systematically investigate the influence of between-individual factors on the predictive and explanatory ability of the TPB. Indeed, experimental *n-of-1* trials employing theory-based behaviour change techniques could investigate the potentially causal pathways in the TPB constructs at different levels i.e., within and between individuals. The results of trials such as these would be invaluable to the development of a cumulative evidence base for behaviour change theories and associated behaviour change techniques.

It is also worth highlighting the potential limitation for these studies associated with employing single item measures of each TPB construct rather than combining the scores from multiple measures. The reliability and validity of single-item measures has been discussed in the TPB literature (Sutton, 1998) and it is generally recommended that multiple items are used to ensure accuracy of measurement (Ajzen, 2006; Francis et al., 2004). However, in order to reduce participant burden associated with a lengthy diary and to enhance response likelihood, it was decided that single-item measures previously used in an *n-of-1* study would be used (see Section 5.4.2). Nevertheless, future work could examine whether the identified relationships between the TPB constructs and behaviour differ depending on whether single or multiple item TPB measures are used.

Chapter 8 - General discussion

8.1 What Was Known Before This Thesis

Several clinical factors have been identified as influencing clinicians' decision regarding prioritisation of patients for total joint replacement (TJR). Patient psychological factors may also be important. Clinical prioritisation tools are used by some clinicians to aid the prioritisation decision.

There is some evidence that pre-operative patient cognitions can predict health outcomes and recovery after surgery. The evidence for the role of pre-operative patient expectations on recovery after TJR is mixed. Studies lack the application of a theoretical framework limiting the development of a cumulative evidence base in this area.

Implicit in the Theory of Planned Behaviour (TPB) is its application *within* individuals; however empirical support for the utility of the TPB to predict and explain health behaviours is dominated by group-based designs investigating between-individual differences.

8.2 What This Thesis Adds

The International Classification of Functioning, Disability and Health (ICF) and the integrated ICF/TPB model can be used as a theoretical framework for understanding clinical decision making. Prioritisation tools measure all aspects of health and functioning as indexed by the ICF. However, prioritisation tools employ measurement items that are open to interpretation, which may negate the aim of such tools which is to standardise prioritisation decisions. When prioritisation tools are not employed, clinicians' decisions are influenced by patient behavioural (disability) and cognitive factors in addition to impairment-based indicators.

The ICF is a suitable theoretical framework to classify pre-operative patient expectations for total hip replacement (THR). However, there is little support for a relationship between patient expectations, as measured in this thesis, and recovery after surgery.

The TPB can predict physical activity (PA) behaviour *within* individuals with osteoarthritis and *within* some healthy individuals. There is some support for the candidate causal pathways in the TPB *within* individuals with and without mobility problems as revealed by individually tailored TPB-based or action planning interventions.

8.3 General Limitations

There are two main limitations to this thesis. The first limitation relates to the wording of the two items that were used to measure pre-operative patient expectations in the 'EUROHIP' cohort (Chapter 4). It is possible that the use of the verb 'do' within the items, shaped patients' thinking towards behavioural expectations explaining the finding that the majority of reported expectations were classified as activity limitations and participation restrictions and few were classified as impairment.

The second main limitation relates to the interventions employed in the *n-of-1* studies. The inconsistent support for PC or SE interventions within individuals suggests that the design of these interventions could have been improved. Similarly, action planning interventions were given to participants either on the basis of their intention being the strongest predictor of their chosen PA behaviour during the baseline or by default.

However, the strength of the participants' intention for the chosen PA was not taken into consideration, which may mean that participants were not sufficiently motivated to support an action planning intervention. Further, a related limitation is that the TPB measures used in the *n-of-1* studies were not individually tailored to each participant. Therefore, there may have been a lack of measurement sensitivity within the items employed to detect within-individual behaviour change.

8.4 Implications and Future Work

This thesis employed the theoretical frameworks of the ICF, the TPB and the integrated ICF/TPB model to investigate disability and PA behaviours. The ICF is an internationally accepted framework of health and functioning that recognises the disability associated with a particular health condition such as osteoarthritis. Disability can be conceptualised as behaviour, therefore the ICF components of 'activity limitation' and 'participation restriction' can reflect behavioural symptoms commonly reported by individuals with osteoarthritis.

Evidence from the studies investigating clinicians' prioritisation of patients for TJR suggests that with and without the use of prioritisation tools, clinicians' assign priority on the basis of a patient's impairment and the behavioural symptoms of 'activity limitation' and 'participation restriction'. However, when clinicians are also asked to consider the relative importance of a patient psychological factor, taken from the integrated ICF/TPB model, namely patient motivation, the prioritisation decision is made affording more importance to this psychological factor than either a patient's pain (impairment) or disability (activity limitation). None of the items comprising the prioritisation tools (Chapter 2) refer to patient psychological factors; therefore, it would seem that the tools adopt a reductionist view of the indications for TJR, focusing on clinical and disability-related behavioural factors and ignoring potentially important psychological indicators. Replication of the conjoint study (Chapter 3) is needed to validate the finding that patient motivation influences the prioritisation decision. Furthermore, it would be useful to replicate the study with a sample of referring physicians rather than orthopaedic surgeons to investigate whether the prioritisation decision differs by health professional.

Analyses of the ‘EUROHIP’ data (Chapter 4) did not find strong evidence in support of a relationship between pre-operative patient expectations and recovery after THR. However, the measure of patient expectations in this dataset was problematic. Future studies should elicit patient expectations in a manner that does not induce response bias towards a particular type of expectation (i.e., in this case, towards activity-based expectations). In addition, the observation that a clinician’s decision making around TJR appears to be influenced by their perception of a patient’s motivation raises the interesting question of whether there is a relationship between patient motivation and recovery from surgery after TJR. Identification of a positive relationship between higher patient motivation and better recovery after surgery would suggest that prioritisation tools, such as those operationalised in this thesis, would benefit from being revised to incorporate a measure of patient motivation.

The integrated ICF/TPB model rather than the ICF framework alone is a more suitable theoretical framework to employ in future work investigating clinical prioritisation and health outcomes after TJR because it recognises the potential role for psychological constructs. ICF measurement categories for the component of ‘personal factors’, defined as contextual factors that relate to the individual (World Health Organization, 2001), are yet to be agreed. However, the integrated ICF/TPB model advances understanding of the relationship between impairment and activity by the inclusion of psychological constructs of the TPB; these constructs are conceptually compatible with what the ICF defines as ‘personal factors’. For example, Jelsma (2009) notes that patient factors such as motivation, confidence and optimism should be coded as ‘personal factors’, which are not dissimilar to intention and control cognitions in the TPB.

Employing TPB variables to further define the ICF component of ‘personal factors’, has two advantages. First, the current evidence base for the TPB constructs provides a valuable source of empirical support for the role of psychological constructs in predicting and explaining health behaviour. This existing evidence base could function as the foundation from which the ‘personal factors’ categories can be developed further. Second, there is published guidance on developing reliable and valid items to measure TPB and other psychological constructs (Ajzen, 2006; Francis et al., 2004). Therefore, unlike the current problem with the ICF core measurement sets, which identify *what* to measure but not *how* to measure it, TPB derived personal factors will include guidance on *how* to measure those personal factors.

Discussion and research focusing on whether the ICF constructs of *activity limitation* and *participation restriction* are conceptually distinct or combined (Jette et al., 2007; Jette et al., 2003) and whether they can be measured with discriminant validity (Dixon et al., 2007; Dixon et al., 2008a; Pollard et al., 2009) is ongoing. The findings of this thesis add to this debate. The method of discriminant content validation employed in Chapter 2 found that at least one item from each of the prioritisation tools was a discriminant measure of *activity limitation* and at least one item from each tool was a discriminant measure of *participation restriction*. This finding supports the view that the constructs are conceptually distinct. However, both prioritisation tools also contained items that were mixed measures of the combined construct of *activity limitation & participation restriction*. This finding emphasises the problem for the development of *pure* measures of each of the ICF constructs without measurement confound (Pollard et al., 2006).

Illustrated in the studies reported in Chapters 6 and 7, and in previous research (Crane et al., 2003; Jones et al., 1989), and recognised in the MRC Complex Interventions Guidance (Craig et al., 2008), is the ability to test hypotheses and theory *within* individuals using an *n-of-1* design. Information regarding the specific times at which an individual is more likely to engage in a behaviour can have significant implications for the design and implementation of a behaviour change intervention at the individual level and wider population level. For example, at the individual level, a TPB-based PA intervention can be designed in response to an individual's baseline data, as was conducted in the *n-of-1* studies reported in this thesis. At the population level, data indicating that within-individual variability in TPB cognitions can predict PA challenges the currently dominating group design approach, which by nature of its between-individual analyses assumes within-individual stability of cognitions and ignores the possibility of within-individual variability. Instead of implementing a PA intervention to target individuals with weaker cognitions, an intervention could be designed to target individuals at times when the conditions would suggest that their cognitions are likely to be weaker. For example, in individuals with osteoarthritis, this may be at times when symptoms such as pain are worse, whereas in healthy individuals, this may be at times when actual or perceived barriers to PA are high, such as barriers associated with work or family commitments.

The overarching theory applied in these series of experimental *n-of-1* studies was the TPB and the effect of the interventions was investigated in relation to changes in TPB constructs and behaviour. It is recognised that an unmeasured moderating variable may have also influenced the effect of the intervention on the TPB construct and/or behaviour. For example, even if a SE intervention was found to increase SE and behaviour within an individual, it is possible that the individual's mood (e.g., stress, anxiety or depression) may have also contributed to the effect of the intervention. That is, even though mood was not directly targeted nor hypothesised to change as a result of the intervention, a concurrent change in mood may have occurred, which in turn may have contributed to the intervention's effectiveness. In order to be more confident that the increase in SE produced the increase in behaviour, a reverse casual design could be used (Barlow et al., 2009). This design would aim to reverse the effect of the intervention by decreasing SE and observing whether behaviour also decreased, which would suggest that it is the change in SE that produced the change in behaviour. It would also be useful to measure potential moderating variables to ascertain the influence they may have on the effect of an intervention.

Other future work could consider employing qualitative methods alongside the *n-of-1* methodology. For example, 'think aloud' techniques have been employed in studies of the TPB to explore the meaning of the TPB items for the individual (Darker & French, 2009). Within the *n-of-1* methodology it would be useful if the individual verbalised their thoughts whilst filling in the diary to help to elucidate how the individual interprets each item. For example, the intention item used in the series of *n-of-1* studies reported in this thesis was: 'To what extent do you intend to X between now and the next time you fill in the diary?' Intention may have different components, which may include an individual's *intent* to do a behaviour and an individual's *desire* or *want* to do a behaviour. Using a ‘think-aloud’ technique would allow the precise meaning of the intention item for an individual to be identified, which would facilitate interpretation of the data showing the relationships between TPB constructs and behaviour.

An alternative qualitative method which could be used alongside the *n-of-1* approach is focus groups. The design of the behaviour change intervention studies used in this thesis followed a *top-down* approach whereby interventions that were known to be effective in the wider population were identified and then applied at the individual level. Adopting a *bottom-up* approach, a focus group with the sample population could explore and identify which intervention features would be acceptable and appropriate for the target population, and then design the intervention accordingly. If a focus group had been conducted prior to conducting the series of *n-of-1* studies with the older population in Chapter 6, then the issue of whether action planning is an acceptable technique for a retired population, would perhaps have been identified and modified accordingly.

8.5 Anecdotal Observations in the *n-of-1* Studies

8.5.1 Participant Burden and Feedback on Participation

The *n-of-1* methodology necessitates frequent and repeated observations in participants over a period of time. The current studies required twice daily diary entries for a period of 12 weeks, thus, the intensive involvement and concurrent potential burden on participants in an *n-of-1* study is high, which may increase the likelihood of participant attrition and missing diary entries (Stone & Shiffman, 2002). However, in the study reported in Chapter 6, no participants dropped out and compliance to diary completion was high with at least 91% of possible diary entries being filled in. Furthermore, in the study reported in Chapter 7, only one of seven participants dropped out and the rest completed the study with at least 96% of possible diary entries being filled in. This suggests that participant burden was acceptable in these studies. End of study debriefing elicited participants' experience of participation and the general evaluation of participation was positive. The handheld diaries were well received and participants reported that filling in the diary was easy (Chapter 6). The option to use either the web-based or paper diary depending on personal preference and internet accessibility was viewed positively (Chapter 7). When asked whether participants would have considered continuing filling in the diary for more weeks beyond the 12-week protocol, three of four of the individuals with osteoarthritis answered affirmatively (Chapter 6), whilst only one of the six healthy individuals answered affirmatively (Chapter 7). This discrepancy indicates that the acceptable length for participants to take part in an *n-of-1* study is finite and may depend on some characteristic of the population sample; for example, perhaps retired individuals like those in Chapter 6 are more willing to take part for longer periods.

8.5.2 Self-Monitoring Effects and Active Components of the Behaviour Change Interventions

The *n-of-1* studies reported in this thesis used daily diary methods to collect within-individual assessments over time. Within-individual analyses of variability in measures over time and the relationships between measures, such as between cognitions and behaviour, can be conducted. In addition to statistical analyses, intrinsic to daily diary methods is the possibility that participants will engage in self-monitoring of their feelings, cognitions and behaviour. This self-monitoring can lead to self-regulatory behaviour. For example, one individual with osteoarthritis (Chapter 6) said that as a result of filling in the diary he had noticed that he experienced more pain late in the evenings and at night so had altered the time he took his analgesic medication to reflect this. Another individual with osteoarthritis (Chapter 6) reported trying to identify whether walking resulted in more or less pain, and had concluded that walking did not seem to affect pain; therefore, she would not avoid walking at times when the pain was worse as she had previously done. One healthy individual (Chapter 7) said that if at the end of a day they reported a lot less steps than normal, then they would make an extra effort to walk more the next day. These examples illustrate the potentially powerful effects of self-monitoring as a behaviour change technique.

Self-monitoring has been identified as an effective component of health behaviour interventions, including PA interventions, by a large meta-regression of interventions (Michie et al., 2009). Intrinsic to daily diary methods is the likelihood that participants will self-monitor. The *n-of-1* studies reported in this thesis employed diary methods; therefore, it is likely that by virtue of filling in the diary, participants engaged in some self-monitoring during the 12-week study period. Further, self-monitoring was not targeted by the intervention given at study week 6 (i.e., the intervention was not designed to change participants' self-monitoring). Therefore, the increase in PA behaviour identified in some of the participants in the *n-of-1* studies cannot be directly attributed to self-monitoring.

However, interventions that combine self-monitoring with another self-regulatory technique have been found to be more effective than interventions that do not include self-monitoring and any other technique (Michie et al., 2009). Thus, it is possible that even though self-monitoring was a potential factor influencing a participant's behaviour throughout the 12-week study rather than a specific component of the intervention per se, the effectiveness of an intervention targeting PC or SE, for example, may depend on the participant concurrently engaging in self-monitoring. Further, as discussed in Chapters 6 and 7, in addition to the main behaviour change techniques employed in the interventions (i.e., the TPB-based intervention targeting either perceived controllability or self-efficacy, or the action planning intervention), the intervention protocol also included a feedback component where the participant was told what influenced whether he/she engaged in their chosen PA behaviour during the first 6 weeks of the study. Therefore, it is acknowledged that this feedback component may have contributed to the intervention effects. The ongoing work to develop and refine a comprehensive taxonomy of effective behaviour change techniques which map to theoretical constructs (Abraham & Michie, 2008; Michie et al., 2008; Michie et al., 2009) will mean that future *n-of-1* studies will be able to easily select behaviour change techniques on the basis of their theoretical and empirical support. Furthermore, this will allow specific techniques to be tested in conjunction with self-monitoring within individuals.

REFERENCES

Abell, J. E., Hootman, J. M., Zack, M. M., Moriarty, D., & Helmick, C. G. (2005). Physical activity and health related quality of life among people with arthritis. *Journal of Epidemiology and Community Health, 59,* 380-385.

Abraham, C. & Michie, S. (2008). A taxonomy of behavior change techniques used in interventions. *Health Psychology, 27,* 379-387.

Affleck, G., Tennen, H., Urrows, S., & Higgins, P. (1991). Individual differences in the day-to-day experience of chronic pain: a prospective daily study of rheumatoid arthritis patients. *Health Psychology, 10,* 419-426.

Affleck, G., Zautra, A., Tennen, H., & Armeli, S. (1999). Multilevel daily process designs for consulting and clinical psychology: a preface for the perplexed. *Journal of Consulting and Clinical Psychology, 67,* 746-754.

Ainsworth, B. E., Bassett, D. R., Strath, S. J., Swartz, A. M., O'Brien, W. L., Thompson, R. W. et al. (2000). Comparison of three methods for measuring the time spent in physical activity. *Medicine & Science in Sports & Exercise, 32,* S457-S464.

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes, 50,* 179-211.

Ajzen, I. (2002). Perceived behavioral control, self-efficacy, locus of control, and the theory of planned behavior. *Journal of Applied Social Psychology, 32,* 1-20.

Ajzen, I. (2006). *Constructing a TPB questionnaire: conceptual and methodological considerations*. Retrieved March 20, 8 A.D., from http://people.umass.edu/aizen/

Allen, K. D., Coffman, C. J., Golightly, Y. M., Stechuchak, K. M., & Keefe, F. J. (2009). Daily pain variations among patients with hand, hip, and knee osteoarthritis. *Osteoarthritis and Cartilage, 17,* 1275-1282.

American Society of Anesthesiology (1963). New classification of physical status. *Anesthesiology, 24,* 111.

Armitage, C. J. & Conner, M. (2001). Efficacy of the theory of planned behaviour: a meta-analytic review. *British Journal of Social Psychology, 40,* 471-499.

Arnett, G., Hadorn, D. C., & Steering Committee of the Western Canada Waiting List Project (2003). Developing priority criteria for hip and knee replacement: results from the Western Canada Waiting List Project. *Journal of Applied Social Psychology, 46,* 290-296.

Arthritis Care (2004). *OA nation: the most comprehensive UK report of people with osteoarthritis.* London: Arthritis Care.

Asenlof, P., Denison, E., & Lindberg, P. (2005). Individually tailored treatment targeting motor behavior, cognition and disability: 2 experimental single-case studies of patients with recurrent and persistent musculoskeletal pain in primary health care. *Physical Therapy, 85,* 1061-1077.

Badley, E. M. & Crotty, M. (1995). An international comparison of the estimated effect of the aging population on the major cause of disablement, musculoskeletal disorders. *The Journal of Rheumatology, 22,* 1934-1940.

Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychological Review, 84,* 191-215.

Bandura, A. (1992a). Exercise of personal agency through the self-efficacy mechanism. In R.Schwarzer (Ed.), *Self-efficacy: thought control of action* (pp. 3-38). Washington DC: Hemisphere Publishing Corp.

Bandura, A. (1992b). On rectifying the comparative anatomy of perceived control: comments on "cognates of personal control". *Applied & Preventive Psychology, 1,* 121-126.

Bandura, A. (1997). *Self-efficacy: the exercise of control*. New York: Freeman.

Barlow, D. H., Nock, M. K., & Hersen, M. (2009). *Single case experimental designs: strategies for studying behavior change*. Boston: Pearson Eduction Inc.

Bellamy, N., Watson Buchanan, W., Goldsmith, C. H., Campbell, J., & Stitt, L. W. (1988). Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. *The Journal of Rheumatology, 15,* 1833-1840.

Bickenbach, J. E., Chatterji, S., Badley, E. M., & Ustun, T. B. (1999). Models of disablement, universalism and the International Classification of Impairments, Disabilities and Handicaps. *Social Science & Medicine, 48,* 1173-1187.

Birrell, F., Afzal, C., Nahit, E., Lunt, M., Macfarlane, G. J., Cooper, C. et al. (2002). Predictors of hip joint replacement in new attenders in primary care with hip pain. *British Journal of General Practice, 53,* 26-30.

Birrell, F., Johnell, O., & Silman, A. (1999). Projecting the need for hip replacement over the next three decades: influence of changing demography and threshold for surgery. *Annals of the Rheumatic Diseases, 58,* 569-572.

Bonetti, D. & Johnston, M. (2008). Perceived control predicting the recovery of individual-specific walking behaviours following stroke: testing psychological models and constructs. *British Journal of Health Psychology, 13,* 463-478.

Boonen, A., Stucki, G., Maksymowych, W., Rat, A. C., Escorpizo, R., Boers, M. et al. (2009). The OMERACT-ICF reference group: integrating the ICF into the OMERACT process: opportunities and challenges. *The Journal of Rheumatology, 36,* 2057-2060.

Booth, F. W., Gordon, S. E., Carlson, C. J., & Hamilton, M. T. (2000). Waging war on modern chronic diseases: primary prevention through exercise biology. *Journal of Applied Physiology, 88,* 774-787.

Boult, C., Altmann, M., Gilbertson, D., Yu, C., & Kane, R. L. (1996). Decreasing disability in the 21st century: the future effects of controlling six fatal and nonfatal conditions. *American Journal of Public Health, 86,* 1388-1393.

Brazier, J. E., Harper, R., Jones, N. M. B., O'Cathain, A., Thomas, K. J., Usherwood, T. et al. (1992). Validating the SF-36 health survey questionnaire: new outcome measure for primary care. *British Medical Journal, 305,* 160-164.

British Heart Foundation (2009). *Think fit! pedometer challenge: a guide for participants and team captains.* London: British Heart Foundation.

Brockow, T., Cieza, A., Kuhlow, H., Sigl, T., Franke, T., Harder, M. et al. (2004). Identifying the concepts contained in outcome measures of clinical trials on musculoskeletal disorders and chronic widespread pain using the International Classification of Functioning, Disability and Health as a reference. *Journal of Rehabilitation Medicine, 36,* 30-36.

Brown, D. S., Finkelstein, E. A., Brown, D. R., Buchner, D. M., & Johnson, F. R. (2009). Estimating older adults' preferences for walking programes via conjoint analysis. *American Journal of Preventive Medicine, 36,* 201-207.

Buckley, J. P., Sim, J., Eston, R. G., Hession, R., & Fox, R. (2004). Reliability and validity of measures taken during the Chester step test to predict aerobic power and to prescribe aerobic exercise. *British Journal of Sports Medicine, 38,* 197-205.

Canadian Society for Exercise Physiology (2002). *Physical Activity Readiness Questionnaire*. Retrieved February 2, 2008, from http://www.csep.ca/CMFiles/publications/parq/par-q.pdf

Canavan, A. G. M. (1994). Single-case methodology in clinical neuropsychology. In S.J.E.Lindsay & G. E. Powell (Eds.), *The handbook of clinical adult psychology* (pp. 765-786). London: Routledge.

Cieza, A., Brockow, T., Ewert, T., Amman, E., Kollerits, B., Chatterji, S. et al. (2002). Linking health-status measurements to the International Classification of Functioning, Disability and Health. *Journal of Rehabilitation Medicine, 34,* 205-210.

Cieza, A., Ewert, T., Ustun, T. B., Chatterji, S., Kostanjsek, N., & Stucki, A. (2004). Development of ICF core sets for patients with chronic conditions. *Journal of Rehabilitation Medicine, 36,* 9-11.

Cieza, A., Geyh, S., Chatterji, S., Kostanjsek, N., Ustun, T. B., & Stucki, G. (2005). ICF linking rules: an update based on lessons learned. *Journal of Rehabilitation Medicine, 37,* 212-218.

Cohen, J. (1988). *Statistical power for the behavioral sciences*. New Jersey: Lawrence Erlbaum.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112,* 155-159.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression / correlation analysis for the behavioral sciences*. New Jersey: Lawrence Erlbaum Associates.

Conner, M., Sheeran, P., Norman, P., & Armitage, C. J. (2000). Temporal stability as a moderator of relationships in the theory of planned behaviour. *British Journal of Social Psychology, 39,* 469-493.

Conner-Spady, B. L., Estey, A., Arnett, G., Ness, K., McGurran, J. J., Bear, R. et al. (2004). Prioritization of patients on waiting lists for hip and knee replacement: validation of a priority criteria tool. *International Journal of Technology Assessment in Health Care, 20,* 509-515.

Conner-Spady, B. L., Estey, A., Arnett, G., Ness, K., McGurran, J. J., Bear, R. et al. (2005). Determinants of patient and surgeon perspectives on maximum acceptable waiting times for hip and knee arthroplasty. *Journal of Health Services Research & Policy, 10,* 84-90.

Council of Europe Committee for the Development of Sport and UKK Institute for Health Promotion Research (1995). *Eurofit for adults: assessment of health-related fitness*. Tampere: Council of Europe.

Cowell, I. M. & Phillips, D. R. (2002). Effectiveness of manipulative physiotherapy for the treatment of a neurogenic cervicobrachial pain syndrome: a single case study - experimental design. *Manual Therapy, 7,* 31-38.

Coyte, P. C., Hawker, G., Croxford, R., Attard, C., & Wright, J. G. (1996). Variation in rheumatologists' and familt physicians' perceptions of the indications for and outcomes of knee replacement surgery. *The Journal of Rheumatology, 23,* 730-738.

Craig, P., Dieppe, P., Macintyre, S., Michie, S., Nazareth, I., & Petticrew, M. (2008). *Developing and evaluating complex interventions: new guidance*. Retrieved February 2, 2009, from www.mrc.ac.uk/complexinterventionsguidance

Crane, C., Martin, M., Johnston, D., & Goodwin, G. M. (2003). Does depression influence symptom severity in irritable bowel syndrome? Case study of a patient with irritable bowel syndrome and bipolar disorder. *Psychosomatic Medicine, 65,* 919-923.

Darker, C. D. & French, D. P. (2009). What sense do people make of a theory of planned behaviour questionnaire? A think-aloud study.
Ref Type: Unpublished Work

de Bruin, E. D., Hartmann, A., Uebelhart, D., Murer, K., & Zijlstra, W. (2008). Wearable systems for monitoring mobility-related activities in older people: a systematic review. *Clinical Rehabilitation, 22,* 878.

de Coster, C., McMillan, S., Brant, R., McGurran, J., Noseworthy, T., & Primary Care Panel of the Western Canada Waiting List Project (2007). The Western Canada

Waiting List Project: development of a priority referral score for hip and knee arthroplasty. *Journal of Evaluation in Clinical Practice, 13,* 192-197.

de Groot, K. I., Boeke, S., & Passchier, J. (1999). Preoperative expectations of pain and recovery in relation to postoperative disappointment in patients undergoing lumbar surgery. *Medical Care, 37,* 149-156.

Deberg, M., Dubuc, J.-E., Labasse, A., Sanchez, C., Quettier, E., Bosseloir, A. et al. (2008). One-year follow-up pf Coll2-1, Coll2-1NO2 and myeloperoxidase serum levels in osteoarthritis patients after hip or knee replacement. *Annals of the Rheumatic Diseases, 67,* 168-174.

Dekker, J., Boot, B., van der Woude, L. H. V., & Bijlsma, J. W. J. (1992). Pain and disability in osteoarthritis: a review of biobehavioral mechanisms. *Journal of Behavioral Medicine, 15,* 189-214.

Department of Health (2001). *National service framework for older people.* London: Crown Copyright.

Department of Health (2004a). *At least five a week: evidence on the impact of physical activity and its relationship to health. A report from the Chief Medical Officer.* London: Crown Copyright.

Department of Health (2004b). *Choosing health, choosing activity: a consultation on how to increase physical activity.* London: Crown Copyright.

Department of Health (2006). *The musculoskeletal services framework. a joint responsibility: doing it differently.* London: Crown Copyright.

Department of Health (2009). *NHS 2010-2015: from good to great. preventative, people-centred, productive.* London: Crown Copyright.

Derrett, S., Paul, C., Herbison, P., & Williams, H. (2002). Evaluation of explicit prioritisation for elective surgery: a prospective study. *Journal of Health Services Research, 7,* 14-22.

Dieppe, P., Basler, H.-D., Chard, J., Croft, P., Dixon, J., Hurley, M. et al. (1999). Knee replacement surgery for osteoarthritis: effectiveness, practice variations, indications and possible determinants of utilization. *Rheumatology, 38,* 73-83.

Dieppe, P., Judge, A., Williams, S., Ikwueke, I., Guenther, K.-P., Floeren, M. et al. (2009). Variations in the pre-operative status of patients coming to primary hip replacement for osteoarthritis in european orthopaedic centres. *BMC Musculoskeletal Disorders, 10,* 19.

Dixon, D. & Johnston, M. (2008). Cognitive representations of disability behaviours in people with mobility limitations: consistency with theoretical constructs. *Disability and Rehabilitation, 30,* 126-133.

Dixon, D., Johnston, M., McQueen, M. M., & Court-Brown, C. (2008a). The disabilities of the arm, shoulder and hand questionnaire (DASH) can measure the impairment, activity limitations and participation restriction constructs from the International Classification of Functioning, Disability and Health (ICF). *BMC Musculoskeletal Disorders, 9,* 114.

Dixon, D., Johnston, M., Rowley, D., & Pollard, B. (2008b). Using the ICF and psychological models of behavior to predict mobility limitations. *Rehabilitation Psychology, 53,* 191-200.

Dixon, D., Pollard, B., & Johnston, M. (2007). What does the chronic pain grade questionnaire measure? *Pain, 130,* 249-253.

Dixon, K. E., Keefe, F. J., Scipio, C. D., Perri, L. M., & Abernethy, A. P. (2007). Psychological interventions for arthritis pain management in adults: a meta-analysis. *Health Psychology, 26,* 241-250.

Dolan, P. (1997). Modeling valuations for EuroQol health states. *Medical Care, 35,* 1095-1108.

Dolin, S. J., de C Williams, A. C., Ashford, N., George, J., Pereira, L., & Perello, A. (2003). Factors affecting medical decision-making in patients with osteoarthritis of the hip: allocation of surgical priority. *Disability and Rehabilitation, 25,* 771-777.

Dowell, A. C. & Biran, L. A. (1990). Problems in using the hospital anxiety and depression scale for screening patients in general practice. *British Journal of General Practice, 40,* 27-28.

Dreinhofer, K., Stucki, G., Ewert, T., Huber, E., Ebenbichler, G., Gutenbrunner, C. et al. (2004). ICF core sets for osteoarthritis. *Journal of Rehabilitation Medicine, 36,* 75-80.

Dreinhofer, K. E., Dieppe, P., Sturmer, T., Grober-Gratz, D., Floren, M., Gunther, K.-P. et al. (2006). Indications for total hip replacement: comparison of assessments of orthopaedic surgeons and referring physicians. *Annals of the Rheumatic Diseases, 65,* 1346-1350.

Dzewaltowski, D. A., Noble, J. M., & Shaw, J. M. (1990). Physical activity participation: social cognitive theory versus the theories of reasoned action and planned behaviour. *Journal of Sport & Exercise Psychology, 12,* 388-405.

Ebinesan, A. D., Sarai, B. S., Walley, G., Bridgman, S., & Maffulli, N. (2006). Total knee arthoplasty: good agreement of clinical severity scores between patients and consultants. *BMC Musculoskeletal Disorders, 7,* 61.

Edwards, R. T. (1999). Points for pain: waiting list priority scoring systems. *British Medical Journal, 318,* 412-414.

Eisler, T., Svensson, O., Tengstrom, A., & Elmstedt, E. (2002). Patient expectation and satisfaction in revision total hip arthroplasty. *The Journal of Arthroplasty, 17,* 457-462.

Escobar, A., Gonzalez, M., Quintana, J. M., Bilbao, A., & Ibanez, B. (2009). Validation of a prioritization tool for patients on the waiting list for total hip and knee replacements. *Journal of Evaluation in Clinical Practice, 15*, 97-102.

Ethgen, O., Bruyere, O., Richy, F., Dardennes, C., & Reginster, J.-Y. (2004). Health-related quality of life in total hip and total knee arthroplasty. *The Journal of Bone & Joint Surgery, 86,* 963-974.

Fife-Shaw, C., Sheeran, P., & Norman, P. (2007). Simulating behaviour change interventions based on the theory of planned behaviour: impacts on intention and action. *British Journal of Social Psychology, 46,* 43-68.

Fisher, K. & Johnston, M. (1996). Experimental manipulation of perceived control and its effect on disability. *Psychology & Health, 11,* 657-669.

Fisher, P. L. & Wells, A. (2008). Metacognitive therapy for obsessive-compulsive disorder: a case series. *Journal of Behavior Therapy and Experimental Psychiatry, 39,* 117-132.

Flood, A. B., Lorence, D. P., Ding, J., McPherson, K., & Black, N. A. (1993). The role of expectations in patients' reports of post-operative outcomes and improvement following therapy. *Medical Care, 31,* 1043-1056.

Focht, B. C., Ewing, V., Gauvin, L., & Rejeski, W. J. (2002). The unique and transient impact of acute exercise on pain perception in older, overweight, or obese adults with knee osteoarthritis. *Annals of Behavioral Medicine, 24,* 201-210.

Fortin, P. R., Clarke, A. E., Joseph, L., Liang, M. H., Tanzer, M., Ferland, D. et al. (1999). Outcomes of total hip and knee replacement. *Arthritis 1, 42,* 1722-1728.

Fraenkel, L., Bodardus, S., & Wittink, D. R. (2001). Understanding patient preferences for the treatment of lupus nephritis with adaptive conjoint analysis. *Medical Care, 39,* 1203-1216.

Francis, J. J., Eccles, M. P., Johnston, M., Walker, A., Grimshaw, J., Foy, R. et al. (2004). *Constructing questionnaires based on the theory of planned behaviour: a manual for health services researchers*. Retrieved March 20, 8 A.D., from www.rebeqi.org

Frankel, S., Eachus, J., Pearson, N., Greenwood, R., Chan, P., Peters, T. J. et al. (1999). Population requirement for primary hip-replacement surgery: a cross-sectional study. *The Lancet, 353,* 1304-1309.

Fricker, J. (1999). BMA proposes strategy to reformulate waiting lists. *British Medical Journal, 318,* 78.

Fries, J. F. (1996). Physical activity, the compression of morbidity, and the health of the elderly. *Journal of Royal Society of Medicine, 89,* 64-68.

Furnham, A. F., Hassomal, A., & McClelland, A. (2002). A cross-cultural investigation of the factors and biases involved in the allocation of scarce medical resources. *Journal of Health Psychology, 7,* 381-391.

Glozier, N., Groom, G., & Prince, M. (2004). Patient psychological characteristics have minimal influence on surgical prioritization. *Psychosomatic Medicine, 66,* 251-257.

Godin, G. & Kok, G. (1996). The theory of planned behavior: a review of its applications to health-related behaviors. *American Journal of Health Promotion, 11,* 87-98.

Godin, G. & Shephard, R. J. (1985). A simple method to assess exercise behavior in the community. *Canadian Journal of Applied Sport Science, 10,* 141-146.

Gollwitzer, P. M. (1993). Goal achievement: the role of intentions. *European Review of Social Psychology, 4,* 141-185.

Gollwitzer, P. M. (1999). Implementation intentions: strong effects of simple plans. *American Psychologist, 54,* 493-503.

Gollwitzer, P. M. & Sheeran, P. (2006). Implementation intentions and goal achievement: a meta-analysis of effects and processes. *Advances in Experimental Social Psychology, 38,* 69-119.

Gooberman-Hill, R., Woolhead, G., MacKichan, F., Ayis, S., Williams, S., & Dieppe, S. (2007). Assessing chronic joint pain: lessons from a focus group study. *Arthritis & Rheumatism-Arthritis Care & Research, 57,* 666-671.

Gratton, L., Povey, R., & Clark-Carter, D. (2007). Promoting children's fruit and vegetable consumption: interventions using the theory of planned behaviour as a framework. *British Journal of Health Psychology, 12,* 639-650.

Green, P. E. & Wind, Y. (1975). New ways to measure consumers' judgments. *Harvard Business Review, 53,* 107-117.

Gunnarsdottir, A. I. & Kinnear, M. (2005). Factors that influence prescribers in their selection and use of COX-2 selective inhibitors as opposed to non-selective NSAIDs. *Pharmacy World & Science, 27,* 316-320.

Guyatt, G. H., Heyting, A., Jaeschke, R., Keller, J., Adachi, J. D., & Roberts, R. S. (1990). N of 1 randomized trials for investigating new drugs. *Controlled Clinical Trials, 11,* 88-100.

Gwaltney, C. J., Shields, A. L., & Shiffman, S. (2008). Equivalence of electronic and paper-and-pencil administration of patient-reported outcome measures: a meta-analytic review. *Value in Health, 11,* 322-333.

Hadorn, D. C. & Holmes, A. C. (1997). The New Zealand priority criteria project: part 1: overview. *British Medical Journal, 314,* 131-134.

Hagger, M. S., Chatzisarantis, N. L. D., & Biddle, S. J. H. (2002). A meta-analytic review of the theories of reasoned action and planned behavior in physical activity:

predictive validity and the contribution of additional variables. *Journal of Sport & Exercise Psychology, 24,* 3-32.

Hamaker, E. L., Dolan, C. V., & Molenaar, P. C. M. (2005). Statistical modeling of the individual: rationale and application of multivariate stationary time series analysis. *Multivariate Behavioral Research, 40,* 207-233.

Hardeman, W., Johnston, M., Johnston, D., Bonetti, D., Wareham, N., & Kinmonth, A. L. (2002). Application of the theory of planned behaviour in behaviour change interventions: a systematic review. *Psychology & Health, 17,* 123-158.

Harry, L. E., Nolan, J. F., Elender, F., & Lewis, J. C. (2000). Who gets priority? Waiting list assessment using a scoring system. *Annals of the Royal College of Surgeons of England, 82,* 186-188.

Haskell, W. L., Lee, I.-M., Pate, R. R., Powell, K. E., Blair, S. N., Franklin, B. A. et al. (2007). Physical activity and public health: updated recommendation for adults from the American College of Sports Medicine and the American Heart Association. *Circulation, 116,* 1081-1093.

Hawker, G., Melfi, C., Paul, J., Green, R., & Bombardier, C. (1995). Comparison of a generic (SF-36) and a disease specific (WOMAC) instrument in the measurement of outcomes after knee replacement surgery. *The Journal of Rheumatology, 22,* 1193-1196.

Hawker, G. A., Stewart, L., French, M. R., Cibere, J., Jordan, J. M., March, L. et al. (2008). Understanding the pain experience in hip and knee osteoarthritis- an OARSI/OMERACT initiative. *Osteoarthritis and Cartilage, 16,* 415-422.

Health Policy and Economic Research Unit (1998). *Waiting list prioritisation scoring systems: discussion paper no.6.* London: British Medical Association.

Henn, R. F., Tashjian, R. Z., Kang, L., & Green, A. (2007). Patients' preoperative expectations predict the outcome of rotator cuff repair. *Journal of Bone and Joint Surgery, 89,* 1913-1919.

Hillsdon, M., Foster, C., Cavill, N., Crombie, H., & Naidoo, B. (2005). *The effectiveness of public health interventions for increasing physical activity among adults: a review of reviews. Evidence briefing (2nd Edition)*. Retrieved July 27, 2008, from http://www.nice.org.uk/p.aspx?o=505281

Hillsdon, M. & Thorogood, M. (1996). A systematic review of physical activity promotion strategies. *British Journal of Sports Medicine, 30,* 84-89.

Hirani, V. & Malbut, K. (2002). *Health Survey for England 2000: disability among older people.* London: Stationary Office.

Hochberg, Y. (1988). A sharper Bonferroni procedure for multiple tests of significance. *Biometrika, 75,* 800-803.

Horne, R. (1997). Representations of medication and treatment: advances in theory and measurement. In K.J.Petrie & J. A. Weinman (Eds.), *Perceptions of health & illness* (pp. 155-188). Amsterdam: Harwood Academic Publishers.

Imrie, R. (2004). Demystifying disability: a review of the International Classification of Functioning, Disability and Health. *Sociology of Health & Illness, 26,* 287-305.

Jackson, C., Lawton, R., Knapp, P., Raynor, D. K., Conner, M., Lowe, C. et al. (2005). Beyond intention: do specific plans increase health behaaviours in patients in primary care? A study of fruit and vegetable consumption. *Social Science & Medicine, 60,* 2383-2391.

Jelsma, J. (2009). Use of the International Classification of Functioning, Disability and Health: a literature survey. *Journal of Rehabilitation Medicine, 41,* 1-12.

Jette, A. M., Haley, S. M., & Kooyoomjian, J. T. (2003). Are the ICF activity and participation dimensions distinct? *Journal of Rehabilitation Medicine, 35,* 145-149.

Jette, A. M., Tao, W., & Haley, S. M. (2007). Blending activity and participation sub-domains of the ICF. *Disability and Rehabilitation, 29,* 1742-1750.

Johnston, M. (1996). Models of disability. *Physiotherapy Theory and Practice, 12,* 131-141.

Johnston, M. (1997). Representations of disability. In K.J.Petrie & J. Weinman (Eds.), *Perceptions of health and illness* (pp. 189-212). Amsterdam: Harwood Academic Publishers.

Johnston, M., Bonetti, D., Joice, S., Pollard, B., Morrison, V., Francis, J. J. et al. (2007). Recovery from disability after stroke as a target for a behavioural intervention: results from a randomized controlled trial. *Disability and Rehabilitation, 29,* 1117-1127.

Johnston, M., Bonetti, D., & Pollard, B. (2002). Disability as behaviour: models of measurement and explanation. In L.Backman & C. V. Hofston (Eds.), *Psychology at the turn of the millennium* (pp. 319-333). Sussex: Psychology Press.

Johnston, M., Dixon, D., Hannaford, P., & Elliott, A. (2009). *Disability in the community: role of chronic pain and illness-related cognitions. Chief Scientist Office final report: C29/1/146.*

Johnston, M., Morrison, V., Macwalter, R., & Partridge, C. (1999). Perceived control, coping and recovery from disability following stroke. *Psychology & Health, 14,* 181-192.

Johnston, M. & Pollard, B. (2001). Consequences of disease: testing the WHO international classification of Impairments, Disabilities and Handicaps (ICIDH) model. *Social Science & Medicine, 53,* 1261-1273.

Johnston, M., Pollard, B., Morrison, V., & Macwalter, R. (2004). Functional limitations and survival following stroke: psychological and clinical predictors of 3-year outcome. *International Journal of Behavioral Medicine, 11,* 187-196.

Jones, C. A., Beaupre, L. A., Johnston, D. W. C., & Suarez-Almazor, M. E. (2007). Total joint arthroplasties: current concepts of patient outcomes after surgery. *Rheumatic Disease Clinics of North America, 33,* 71-86.

Jones, K., Johnston, M., & Speck, P. (1989). Despair felt by the patient and the professional carer: a case study of the use of cognitive behavioural methods. *Palliative Medicine, 3,* 39-46.

Jordan, K. M., Arden, N. K., Doherty, M., Bannwarth, B., Bijlsma, J. W. J., Dieppe, P. et al. (2003). EULAR recommendations 2003: an evidence based approach to the management of knee osteoarthritis: report of a task force of the standing committee for international clinical studies including therapeutic trials (ESCISIT). *Annals of the Rheumatic Diseases, 62,* 1145-1155.

Juni, P., Dieppe, P., Donovan, J., Peters, T., Eachus, J., Pearson, N. et al. (2003). Population requirement for primary knee replacement surgery: a cross-sectional study. *Rheumatology, 42,* 516-521.

Kaplan, R. M. (2000). Two pathways to prevention. *American Psychologist, 55,* 382-396.

Kellar, I. & Abraham, C. (2005). Randomized controlled trial of a brief research-based intervention promoting fruit and vegetable consumption. *British Journal of Health Psychology, 10,* 543-558.

Kellett, N., West, F., & Finlay, A. Y. (2006). Conjoint analysis: a novel, rigorous tool for determining patient preferences for topical antibiotic treatment for acne. A randomised controlled trial. *British Journal of Dermatology, 154,* 524-532.

Kelley, K. & Abraham, C. (2004). RCT of a theory-based intervention promoting healthy eating and physical activity amongst out-patients older than 65 years. *Social Science & Medicine, 59,* 787-797.

Kellgren, J. H. & Lawrence, J. S. (1957). Radiological assessment of osteo-arthritis. *Annals of the Rheumatic Diseases, 16,* 494-502.

Kendall, M. G. (1948). *Rank correlation methods*. London: Griffin.

Kiene, S. M., Tennen, H., & Armeli, S. (2008). Today I'll use a condom, but who knows about tomorrow: a daily process study of variability in predictors of condom use. *Health Psychology, 27,* 463-472.

Kopp, M., Bonatti, H., Haller, C., Rumpold, G., Sollner, W., Holzner, B. et al. (2003). Life satisfaction and active coping style are important predictors of recovery from surgery. *Journal of Psychosomatic Research, 55,* 371-377.

Landis, J. R. & Koch, G. G. (1977). Measurement of observer agreement for categorical data. *Biometrics, 33,* 159-174.

Lau-Walker, M. (2006). A conceptual care model for individualized care approach in cardiac rehabilitation - combining both illness representation and self-efficacy. *British Journal of Health Psychology, 11,* 103-117.

Learmonth, I. D., Young, C., & Rorabeck, C. (2007). The operation of the century: total hip replacement. *Lancet, 370,* 1508-1519.

Leventhal, H., Leventhal, E. A., & Contrada, R. J. (1998). Self-regulation, health, and behavior: a perceptual-cognitive approach. *Psychology & Health, 13,* 717-733.

Lippke, S. & Ziegelmann, J. P. (2008). Theory-based health behavior change: developing, testing, and applying theories for evidence-based interventions. *Applied Psychology: an international review, 57,* 698-716.

Lorig, K., Mazonson, P. D., & Holman, H. R. (1993). Evidence suggesting that health education for self-management in patients with chronic arthritis has sustained health benefits while reducing health care costs. *Arthritis & Rheumatism, 36,* 439-446.

Lorig, K., Selenznick, M., Lubeck, D., Ung, E., Chastain, R. L., & Holman, H. R. (1989). The beneficial outcomes of the arthritis self-management course are not adequately explained by behavior change. *Arthritis & Rheumatism, 32,* 91-95.

Lorig, K., Sobel, D. S., Stewart, A. L., Brown, B. W., Bandura, A., Ritter, P. et al. (1999). Evidence suggesting that a chronic disease self-management program can improve health status while reducing hospitalization: a randomized trial. *Medical Care, 37,* 5-14.

Lorig, K. R., Ritter, P. L., Laurent, D. D., & Fries, J. F. (2004). Long-term randomized controlled trials of tailored-print and small-group arthritis self-management interventions. *Medical Care, 42,* 346-354.

Luce, R. D. & Tukey, J. W. (1964). Simultaneous conjoint measurement - a new type of fundamental measurement. *Journal of Mathematical Psychology, 1,* 1-27.

Luszczynska, A. (2006). An implementation intentions intervention, the use of a planning strategy, and physical activity after myocardial infarction. *Social Science & Medicine, 62,* 900-908.

Luszczynska, A. & Tryburcy, M. (2008). Effects of a self-efficacy intervention on exercise: the moderating role of diabetes and cardiovascular diseases. *Applied Psychology: an international review, 57,* 644-659.

Lutz, G. K., Butzlaff, M. E., Atlas, S. J., Keller, R. B., Singer, D. E., & Deyo, R. A. (1999). The relation between expectations and outcomes in surgery for sciatica. *Journal of General Internal Medicine, 14,* 740-744.

MacCormick, A. D., Collecutt, W. G., & Parry, B. R. (2003). Prioritizing patients for elective surgery: a systematic review. *Australian & New Zealand Journal of Surgery, 73,* 633-642.

MacDonald, C. W., Whitman, J. M., Cleland, J. A., Smith, M., & Hoeksma, H. L. (2006). Clinical outcomes following manual physical therapy and exercise for hip osteoarthritis: a case series. *Journal of Orthopaedic & Sports Physical Therapy, 36,* 588-599.

Mahomed, N. N., Liang, M. H., Cook, E. F., Daltroy, L. H., Fortin, P. R., Fossel, A. H. et al. (2002). The importance of patient expectations in predicting functional outcomes after total joint arthroplasty. *The Journal of Rheumatology, 29,* 1273-1279.

Mahon, J. L., Bourne, R. B., Rorabeck, C. H., Feeny, D. H., Stitt, L., & Webster-Bogaert, S. (2002). Health-related quality of life and mobility of patients awaiting elective total hip arthroplasty: a prospective study. *Canadian Medical Association Journal, 167,* 1115-1121.

Maillefert, J. F., Roy, C., Cadet, C., Nizard, R., Berdah, L., & Ravaud, P. (2008). Factors influencing surgeons' decisions in the indication for total joint replacement in hip osteoarthritis in real life. *Arthritis & Rheumatism-Arthritis Care & Research, 59,* 255-262.

Mancuso, C. A., Ranawat, C. S., Esdaile, J. M., Johanson, N. A., & Charlson, M. E. (1996). Indications for total hip and total knee arthroplasties: results of orthopaedic surveys. *The Journal of Arthroplasty, 11,* 34-46.

Mancuso, C. A., Salvati, E. A., Johanson, N. A., Peterson, M. G. E., & Charlson, M. E. (1997). Patients' expectations and satisfaction with total hip arthroplasty. *The Journal of Arthroplasty, 12,* 387-396.

Mannion, A. F., Kampfen, S., Munzinger, U., & Kramers-de Quervain, I. (2009). The role of patient expectations in predicting outcome after total knee arthroplasty. *Arthritis Research & Therapy, 11,* R139.

March, L., Irwig, L., Schwarz, J., Simpson, J., Chock, C., & Brooks, P. (1994). n of 1 trials comparing a non-steroidal anti-inflammatory drug with paracetamol in osteoarthritis. *British Medical Journal, 309,* 1041-1046.

March, L. M., Cross, M. J., Lapsley, H., Brnabic, A. J. M., Tribe, K. L., Bachmeier, C. J. M. et al. (1999). Outcomes after hip or knee replacement surgery for osteoarthritis. *Medical Journal of Australia, 171,* 235-238.

Marshall, S. J., Levy, S. S., Tudor-Locke, C. E., Kolkhorst, F. W., Wooten, K. M., Ji, M. et al. (2009). Translating physical activity recommendations into a pedometer-based step goal. *American Journal of Preventive Medicine, 36,* 410-415.

Martin, J., Meltzer, H., & Elliott, D. (1988). *The prevalence of disability among adults. OPCS surveys of disability in Great Britain.* London: Her Majesty's Stationary Office.

Mathews, A. & Ridgeway, V. (1981). Personality and surgical recovery: a review. *British Journal of Clinical Psychology, 20,* 243-260.

McGartland Rubio, D., Berg-Weger, M., Tebb, S. S., Lee, E. S., & Rauch, S. (2003). Objectifying content validity: conducting a content validity study in social work research. *Social Work Research, 27,* 94-104.

McGraw, K. O. & Wong, S. P. (1996). Forming inferences about some intraclass corrleation coefficients. *Psychological Methods, 1,* 30-46.

McNair, P. J., Simmonds, M. A., Boocock, M. G., & Larmer, P. J. (2009). Exercise therapy for the management of osteoarthritis of the hip joint: a systematic review. *Arthritis Research & Therapy, 11,* R98.

Messier, S. P., Loeser, R. F., Miller, G. D., Morgan, T. M., Rejeski, W. J., Sevick, M. A. et al. (2004). Exercise and dietary weight loss in overweight and obese older adults with knee osteoarthritis: the arthritis, diet, and activity promotion trial. *Arthritis & Rheumatism, 50,* 1501-1510.

Michie, S. & Abraham, C. (2004). Interventions to change health behaviours: Evidence-based or evidence-inspired? *Psychology & Health, 19,* 29-49.

Michie, S., Abraham, C., Whittington, C., McAteer, J., & Gupta, S. (2009). Effective techniques in healthy eating and physical activity interventions: a meta-regression. *Health Psychology, 28,* 690-701.

Michie, S., Dormandy, E., & Marteau, T. M. (2004). Increasing screening uptake amongst those intending to be screened: the use of action plans. *Patient Education and Counseling, 55,* 218-222.

Michie, S., Johnston, M., Abraham, C., Lawton, R., Parker, D., & Walker, A. (2005). Making psychological theory useful for implementing evidence based practice: a consensus approach. *Quality and Safety in Health Care, 14,* 26-33.

Michie, S., Johnston, M., Francis, J., Hardeman, W., & Eccles, M. P. (2008). From theory to intervention: mapping theoretically derived behavioural determinants to behaviour change techniques. *Applied Psychology: an international review, 57,* 660-680.

Michie, S. & Prestwich, A. (2010). Are interventions theory-based? development of a theory coding scheme. *Health Psychology, 29,* 1-8.

Miller, G. D., Rejeski, W. J., Williamson, J. D., Morgan, T., Sevick, M. A., Loeser, R. F. et al. (2003). The arthritis, diet and activity promotion trial (ADAPT): design, rationale, and baseline results. *Controlled Clinical Trials, 24,* 462-480.

Milne, S., Orbell, S., & Sheeran, P. (2002). Combining motivational and volitional interventions to promote exercise participation: protection motivation theory and implementation intentions. *British Journal of Health Psychology, 7,* 163-184.

Mokdad, A. H., Marks, J. S., Stroup, D. F., & Gerberding, J. L. (2004). Actual causes of death in the united states, 2000. *Journal of American Medical Association, 291,* 1238-1245.

Molenaar, P. C. M. (2004). A manifesto on psychology as idiographic science: bringing the person back into scientific psychology, this time forever. *Measurement, 2,* 210-218.

Molenaar, P. C. M. & Campbell, C. G. (2009). The new person-specific paradigm in psychology. *Current Directions in Psychological Science, 18,* 112-117.

Molloy, G. J., Sniehotta, F., & Johnston, M. (2009). Two alternative models of health behaviour and recovery from activity limitations due to acute injury: a prospective study. *Psychology & Health, 24,* 271-285.

Mondloch, M. V., Cole, D. C., & Frank, J. W. (2001). Does how you do depend on how you think you'll do? A systematic review of the evidence for a relation between patients' recovery expectations and health outcomes. *Canadian Medical Association Journal, 65,* 174-179.

Morley, S. (1994). Single-case methodology in psychological therapy. In S.J.E.Lindsay & G. E. Powell (Eds.), *The handbook of clinical adult psychology* (pp. 722-745). London: Routledge.

Morley, S. & Adams, M. (1991). Graphical analysis of single-case time series data. *British Journal of Clinical Psychology, 30,* 97-115.

Morris, J. N. & Hardman, A. E. (1997). Walking to health. *Sports Medicine, 23,* 306-332.

National Center for Chronic Disease Prevention and Health Promotion: (2007). *Behavioral risk factor surveillance system prevalence and trends data: physical activity 2007*. Retrieved October 30, 9 A.D., from http://www.apps.nccd.cdc.gov/BRFSS/page.asp?yr=2007&state=All#PA

National Health Service Information Centre for health and social care (2009). *Statistics on obesity, physical activity and diet: England February 2009*. Retrieved October 30, 9 A.D., from http://www.ic.nhs.uk/webfiles/publications/opan09/OPAD%20Feb%202009%20finalpdf

National Insitute of Health Consensus Panel (1995). NIH consensus conference: total hip replacement. *The Journal of the American Medical Association, 273,* 1950-1956.

National Institute for Health and Clinical Excellence (2006). *Four commonly used methods to increase physical activity: brief interventions in primary care, exercise referrals schemes, pedometers and community-based exercise programmes for walking and cycling.* London: NICE.

National Institute for Health and Clinical Excellence (2008). *Osteoarthritis: the care and management of osteoarthritis in adults.* London: NICE.

National Joint Registry (2007). *National Joint Registry for England and Wales: 4th Annual Report*. Retrieved June 6, 8 A.D., from http://www.njrcentre.org.uk/statsonline/allsummarymap.php

Naylor, C. D., Williams, J. I., & Ontario panel on hip and knee replacement (1996). Primary hip and knee replacement surgery: Ontario criteria for case selection and surgical priority. *Quality in Health Care, 5,* 20-30.

Nilsdotter, A. K., Toksvig-Larsen, S., & Roos, E. M. (2009). Knee arthroplasty: are patients' expectations fulfilled? *Acta Orthopaedica, 80,* 55-61.

Norman, P. & Conner, M. (2005). The theory of planned behavior and exercise: evidence for the mediating and moderating roles of planning on intention-behavior relationships. *Journal of Sport & Exercise Psychology, 27,* 488-504.

Noseworthy, T. W., McGurran, J. J., Hadorn, D. C., & Steering Committee of the Western Canada Waiting List Project (2003). Waiting for scheduled services in Canada: development of priority-setting scoring systems. *Journal of Evaluation in Clinical Practice, 9,* 23-31.

Ogilvie, D., Foster, C. E., Rothnie, H., Cavill, N., Hamilton, V., Fitzsimons, C. F. et al. (2007). Interventions to promote walking: systematic review. *British Medical Journal, 334,* 1204.

Onghena, P. & Edgington, E. S. (2005). Customization of pain treatments: single-case design and analysis. *Clinical Journal of Pain, 21,* 56-68.

Orbell, S., Espley, A., Johnston, M., & Rowley, D. (1998). Health benefits of joint replacement surgery for patients with osteoarthritis: prospective evaluation using independent assessments in Scotland. *Journal of Epidemiology and Community Health, 52,* 564-570.

Orbell, S., Johnston, M., Rowley, D., Espley, A., & Davey, P. (1998). Cognitive representations of illness and functional and affective adjustment following surgery for osteoarthritis. *Social Science & Medicine, 47,* 93-102.

Orme, B. K. (2006). Sample size issues for conjoint analysis. In B.K.Orme (Ed.), *Getting started with conjoint analysis: strategies for product design and pricing research* (pp. 49-57). Madison: Research Publishers LLC.

Ottenbacher, K. J. (1990). Clinically relevant designs for rehabilitation research: the idiographic model. *American Journal of Physical Medicine & Rehabilitation, 69,* 286-292.

Oudhoff, J. P., Timmermans, D. R. M., Knol, D. L., Bijnen, A. B., & van der Wal, G. (2007). Prioritising patients on surgical waiting lists: a conjoint analysis study on the priority judgements of patients, surgeons, occupational physicians, and general practitioners. *Social Science & Medicine, 64,* 1863-1875.

Perenboom, R. J. M. & Chorus, A. M. J. (2003). Measuring participation according to the international classification of functioning, disability and health (ICF). *Disability and Rehabilitation, 25,* 577-587.

Petrie, K. J., Cameron, L. D., Ellis, C. J., Buick, D., & Weinman, J. (2002). Changing illness perceptions after myocardial infarction: an early intervention randomized controlled trial. *Psychosomatic Medicine, 64,* 580-586.

Petrie, K. J., Weinman, J., Sharpe, N., & Buckley, J. (1996). Role of patients' view of their illness in predicting return to work and functioning after myocardial infarction: longitudinal study. *British Medical Journal, 312,* 1191-1194.

Pollard, B., Dixon, D., Dieppe, P., & Johnston, M. (2009). Measuring the ICF components of impairment, activity limitation and participation restriction: an item analysis using classical test theory and item response theory. *Health and Quality of Life Outcomes, 7,* 41.

Pollard, B., Johnston, M., & Dieppe, P. (2006). What do osteoarthritis health outcome instruments measure? Impairment, activity limitation, or participation restriction? *The Journal of Rheumatology, 33,* 757-763.

Quintana, J. M., Escobar, A., Bilbao, A., Arostegui, I., Lafuente, I., & Vidaurreta, I. (2005). Responsiveness and clinically important differences for the WOMAC and SF-36 after hip joint replacement. *Osteoarthritis and Cartilage, 13,* 1076-1083.

Raley, J. C., Followwill, K. A., Zimet, G. D., & Ault, K. A. (2004). Gynecologists' attitudes regarding human papilloma virus vaccination: a survey of fellows of the american college of obstetricians and gynecologists. *Infectious Diseases in Obstetrics and Gynecology, 12,* 127-133.

Rat, A. C., Guillemin, F., & Pouchot, J. (2008). Mapping the osteoarthritis knee and hip quality of life (OAKHQOL) instrument to the International Classification of Functioning, Disability and Health and comparison to five health status instruments used in osteoarthritis. *Rheumatology, 47*, 1719-1725.

Rissanen, P., Aro, S., Slatis, P., Sintonen, H., & Paavolainen, P. (1995). Health and quality of life before and after hip or knee arthroplasty. *The Journal of Arthroplasty, 10,* 169-175.

Robertson, I. H., Hogg, K., & McMillan, T. M. (1998). Rehabilitation of unilateral neglect: improving function by contralesional limb activation. *Neuropsychological Rehabilitation, 8*, 19-29.

Rosemann, T., Kuehlein, T., Laux, G., & Szecsenyi, J. (2007). Osteoarthritis of the knee and hip: a comparison of factors associated with physical activity. *Clinical Rheumatology, 26,* 1811-1817.

Ross, C. K., Sinacore, J. M., Stiers, W., & Budiman-Mak, E. (1990). The role of expectations and preferences in health care satisfaction of patients with arthritis. *Arthritis Care and Research, 3*, 92-98.

Rutter, D. R., Steadman, L., & Quine, L. (2006). An implementation intentions intervention to increase uptake of mammography. *Annals of Behavioral Medicine, 32,* 127-134.

Ryan, M. (1996). *Using consumer preferences in health care decision making*. London: Office of Health Economics.

Ryan, M. & Farrar, S. (2000). Using conjoint analysis to elicit preferences for health care. *British Medical Journal, 320,* 1530-1533.

Ryan, P. & Lauver, D. R. (2002). The efficacy of tailored interventions. *Journal of Nursing Scholarship, 34,* 331-337.

Salmon, P. (2000). Effects of physical exercise on anxiety, depression, and sensitivity to stress: a unifying theory. *Clinical Psychology Review, 21,* 33-61.

Sawatzky, R., Liu-Ambrose, T., Miller, W. C., & Marra, C. A. (2007). Physical activity as a mediator of the impact of chronic conditions on quality of life in older adults. *Health and Quality of Life Outcomes, 5,* 68.

Scheier, M. F., Matthews, K. A., Owens, J. F., Magovern, G. J., Lefebvre, R. C., Abbott, R. A. et al. (1989). Dispositional optimism and recovery from coronary artery bypass surgery: the beneficial effects on physical and psychological well-being. *Journal of Personality and Social Psychology, 57,* 1024-1040.

Schroder, C. (2008). *Control cognitions and activity limitations in people with disabling conditions.* PhD Thesis: University of Aberdeen.

Schroder, C., Johnston, M., Teunissen, L., Notermans, N., Helders, P., & van Meeteren, N. (2007). Perceived control is a concurrent predictor of activity limitations in patients with chronic idiopathic axonal polyneuropathy. *Archives of Physical Medicine and Rehabilitation, 88,* 63-69.

Scott, E. J., Eves, F. F., French, D. P., & Hoppe, R. (2007). The theory of planned behaviour predicts self-reports of walking, but does not predict step count. *British Journal of Health Psychology, 12,* 601-620.

Shallice, T. (1979). Case study approach in neuropsychological research. *Journal of Clinical Neuropsychology, 1,* 183-211.

Sheeran, P. (2002). Intention-behavior relations: a conceptual and empirical review. *European Review of Social Psychology, 12,* 1-36.

Sheeran, P. & Abraham, C. (2003). Mediator of moderators: temporal stability of intention and the intention-behavior relation. *Personality and Social Psychology Bulletin, 29,* 215.

Sheeran, P., Orbell, S., & Trafimow, D. (1999). Does the temporal stability of behavioral intentions moderate intention-behavior and past behavior-future behavior relations? *Personality and Social Psychology Bulletin, 25,* 724-734.

Sheeran, P. & Silverman, M. (2003). Evaluation of three interventions to promote workplace health and safety: evidence for the utility of implementation intentions. *Social Science & Medicine, 56,* 2153-2163.

Sheeran, P., Webb, T. L., & Gollwitzer, P. M. (2005). The interplay between goal intentions and implementation intentions. *Personality and Social Psychology Bulletin, 31,* 87-98.

Shields, R. K., Enloe, L. J., & Leo, K. C. (1999). Health related quality of life in patients with total hip or knee replacement. *Archives of Physical Medicine and Rehabilitation, 80,* 572-579.

Shiffman, S., Stone, A. A., & Hufford, M. R. (2008). Ecological momentary assessment. *Annual Review of Clinical Psychology, 4,* 1-32.

Singh, J., Cuttler, L., Shin, M., & Neuhauser, D. (1998). Medical decision-making and the patient: understanding preference patterns for growth hormone therapy using conjoint analysis. *Medical Care, 36,* AS31-AS45.

Sniehotta, F. F. (2009). Towards a theory of intentional behaviour change: plans, planning and self-regulation. *British Journal of Health Psychology, 12,* 261-273.

Sniehotta, F. F., Scholz, U., & Schwarzer, R. (2006). Action plans and coping plans for physical exercise: a longitudinal intervention study in cardiac rehabilitation. *British Journal of Health Psychology, 11,* 23-37.

Stanek, E. J., Oates, M. B., McGhan, W. F., Denofrio, D., & Loh, E. (2000). Preferences for treatment outcomes in patient heart failure: symptoms versus survival. *Journal of Cardiac Failure, 6,* 225-232.

Stevens, N. & Sykes, D. (1996). Aerobic fitness testing: an update. *Occupational Health, 48,* 436-438.

Stone, A. A., Kessler, R. C., & Haythornthwaite, J. A. (1991). Measuring daily events and experiences: decisions for the researcher. *Journal of Personality, 59,* 575-607.

Stone, A. A. & Shiffman, S. (2002). Capturing momentary, self-report data: a proposal for reporting guidelines. *Annals of Behavioral Medicine, 24,* 236-243.

Stucki, A., Daansen, P., Fuessl, M., Cieza, A., Huber, E., Atkinson, R. et al. (2004). ICF core sets for obesity. *Journal of Rehabilitation Medicine, 36,* 107-113.

Stucki, G., Cieza, A., Geyh, S., Battistella, L., Lloyd, J., Symmons, D. et al. (2004). ICF core sets for rheumatoid arthritis. *Journal of Rehabilitation Medicine, 36,* 87-93.

Sturmer, T., Dreinhofer, K., Grober-Gratz, D., Brenner, H., Dieppe, P., Puhl, W. et al. (2005). Differences in the views of orthopaedic surgeons and referring practitioners on the determinants of outcome after total hip replacement. *The Journal of Bone & Joint Surgery (British), 87,* 1416-1419.

Sutton, S. (1998). Predicting and explaining intentions and behavior: how well are we doing? *Journal of Applied Social Psychology, 28,* 1317-1338.

Sutton, S. (2003). Testing attitude-behaviour theories using non-experimental data: an examination of some hidden assumptions. *European Review of Social Psychology, 13,* 293-323.

Tabachnick, B. G. & Fidell, L. S. (2001). *Using multivariate statistics*. Boston: Allyn and Bacon.

Taris, T. W. (2000). *A primer in longitudinal data analysis*. London: Sage.

Terry, D. J. & O'Leary, J. E. (1995). The theory of planned behaviour: the effects of perceived behavioural control and self-efficacy. *British Journal of Social Psychology, 34,* 199-220.

The EuroQol Group (1990). EuroQol - a new facilty for the measurement of health-related quality of life. *Health Policy, 16,* 199-208.

The Scottish Government (2008). *The Scottish Health Survey 2008*. Retrieved November 11, 9 A.D., from http://www.scotland.gov.uk/Publications/2009/09/28102003/66

Trafimow, D., Sheeran, P., Conner, M., & Finlay, K. A. (2002). Evidence that perceived behavioural control is a multidimensional construct: perceived control and perceived difficulty. *British Journal of Social Psychology, 41*, 101-121.

Tudor-Locke, C., Williams, J. E., Reis, J. P., & Pluto, D. (2002). Utility of pedometers for assessing physical activity: convergent validity. *Sports Medicine, 32,* 795-808.

Tugwell, P., Bombardier, C., Buchanan, W. W., Goldsmith, C. H., Grace, E., & Hanna, B. (1987). The MACTAR patient preference disability questionnaire- and individualised functional priority approach for assessing improvement in physical disability in clinical trials in rheumatoid arthritis. *The Journal of Rheumatology, 14,* 446-451.

Tuominen, U., Sintonen, H., Hirvonen, J., Seitsalo, S., Paavolainen, P., Lehto, M. et al. (2009). The effect of waiting time on health and quality of life outcomes and cost of medication in hip replacement patients: a randomized clinical trial. *Osteoarthritis and Cartilage, 17,* 1144-1150.

Uhlmann, R. F., Inui, T. S., & Carter, W. B. (1984). Patient requests and expectations. *Medical Care, 22,* 681-685.

Vancouver, J. B. & Kendall, L. N. (2006). When self-efficacy negatively relates to motivation and performance in a learning context. *Journal of Applied Psychology, 91,* 1146-1153.

Vancouver, J. B., Thompson, C. M., & Williams, A. A. (2001). The changing signs in the relationship among self-efficacy, personal goals, and performance. *Journal of Applied Psychology, 86,* 605-620.

Ware, J. E., Snow, K. K., Kosinski, M., & Gandek, B. (1993). *SF-36 health survey: manual and interpretation guide*. Boston, MA: The Health Institute.

Webb, T. L. & Sheeran, P. (2006). Does changing behavioral intentions engender behavior change? A meta-analysis of the experimental evidence. *Psychological Bulletin, 132,* 249-268.

Weigl, M., Cieza, A., Harder, M., Kostanjsek, N., & Stucki, G. (2003). Linking osteoarthritis-specific health-status measures to the International Classification of Functioning, Disability, and Health (ICF). *Osteoarthritis and Cartilage, 11,* 519-523.

Western Canada Waiting List Project (2004). *Western Canada Waiting List Project 2: moving forward; final report.* Calgary: Western Canada Waiting List Project.

Williams, N. H., Hendry, M., France, B., Lewis, R., & Wilkinson, C. (2007). Effectiveness of exercise-referral schemes to promote physical activity in adults: systematic review. *British Journal of General Practice, 57,* 979-986.

World Health Organization (1980). *The International Classification of Impairments, Disabilities and Handicaps.* Geneva: World Health Organization.

World Health Organization (2001). *WHO International Classification of Functioning, Disability and Health: ICF.* Geneva: World Health Organization.

World Health Organization (2003). *Global strategy on diet, physical activity and health.* Geneva: World Health Organization.

World Health Organization (2004). *The global burden of disease: 2004 update.* Geneva: World Health Organization.

Zhang, W., Doherty, M., Arden, N., Bannwarth, B., Bijlsma, J., Gunther, K.-P. et al. (2005). EULAR evidence based recommendations for the management of hip osteoarthritis: report of a task force of the EULAR standing committee for international clinical studies including therapeutics (ESCISIT). *Annals of the Rheumatic Diseases, 64,* 669-681.

Zigmond, A. S. & Snaith, R. P. (1983). The hospital anxiety and depression scale. *Acta Psychiatrica Scandinavica, 67,* 361-370.

APPENDIX 1: New Zealand Priority Criteria for Joint Replacement (NZPC)

Pain

Degree (Patient must be on maximum medical therapy at time of rating)
1. Mild - slight or occasional pain; patient has not altered patterns of activity or work
2. Mild/moderate - moderate or frequent pain; patient has not altered patterns of activity or work
3. Moderate - patient is active, but has had to modify or give up some activities because of pain
4. Moderate/severe - fairly severe pain with substantially limited activities
5. Severe - major pain and serious limitation

Occurrence
6. None or with first steps only
7. Only after long walks (30 minutes)
8. With all walking, mostly day pain
9. Significant, regular night pain

Functional Activity

Time Walked
10. Unlimited
11. 31-60 minutes (eg longer shopping trips to mall)
12. 11-30 minutes (eg gardening, grocery shopping)
13. 2-10 minutes (eg trip to letter box)
14. Less than 2 minutes or indoors only (more or less house bound)
15. Unable to walk
16. Other functional limitations (eg putting on shoes, managing stairs, sitting to standing, sexual activity, recreation or bobbies, walking aids needed) - ranked on 5-point scale from ‘none’ to ‘severe’

Movement and Deformity

17. Pain on examination (Overall results of both active and passive range of motion) - ranked on a 5-point scale from ’none’ to ’severe’
18. Other abnormal findings (Limited to orthopaedic problems eg reduced range of motion, deformity, limp, instability, progressive x ray findings) - ranked on a 5-point scale from ‘none’ to ‘severe’

Other Factors

19. Severity of multiple joint involvement - rated as ‘no, single joint’, ‘yes, each affected joint: mild/moderate in severity’ or ‘yes, severe involvement’ (eg severe rheumatoid arthritis)
20. Ability to work, give care to dependent/s or live independently (Difficulty must be related to affected joint) - rated as ‘not threatened or difficult’ ‘not threatened but more difficult’, ‘threatened but not immediately’ or ‘immediately threatened’

APPENDIX 2: Western Canada Waiting List Criteria for Joint Replacement: WCWL

Pain on motion (e.g. walking, bending)

1. None/mild pain on motion: Patient can move about including walking and bending. They may experience some pain but it does not prevent any activity. They usually do not require pain medication.

2. Moderate pain on motion: Patient can move about including walking and bending. They experience pain most of the time which limits their activities to some degree. For example, patients experience trouble walking up and down stairs or may be uncomfortable standing for long periods of time. They occasionally need pain medication.

3. Severe pain on motion: Patient cannot walk or bend without experiencing pain. The pain restricts their activities in a major way. For example, patients experience pain walking up and down stairs and may not be able to stand for long periods of time. They need pain medication most of the time.

Pain at rest (e.g. while sitting, lying down, or causing sleep disturbances)

4. No pain at rest: Patient does not experience pain when they are sitting or lying down.

5. Mild pain at rest: Patient experiences some pain but it does not disturb their rest when they are sitting or lying down. Pain does not cause sleep disturbance.

6. Moderate pain at rest: Patient experiences pain most of the time which disturbs their rest when they are sitting or lying down. Pain may cause some sleep disturbance and patient may need to take pain medication occasionally.

7. Severe pain at rest: Patient cannot rest in a sitting or lying position without experiencing pain. Patient often gets up in the middle of the night to take pain medication.

Ability to walk without significant pain

8. Over 5 blocks: Patient can walk over 5 blocks without needing to stop due to pain.

9. 1-5 blocks: Patient can walk between 1 and 5 blocks but then must stop due to the pain.

10. <1 block: Patient cannot walk more than 1 block due to pain.

11. Household ambulator: Patient needs a walking aid, such as crutches or a wheel chair, to walk outside the home. Patient uses furniture for support to walk inside the home.

Other functional limitations

12. No functional limitations: Patient can perform all of their daily tasks such as putting on their shoes, climbing stairs, going from sitting to standing, bathing, cooking and recreation activities. They usually do not require assistance. They usually do not need to make minor modifications to complete the task

13. Mild functional limitations: Patient can perform most of their daily tasks such as putting on their shoes, climbing stairs, going from sitting to standing, bathing, cooking and recreation activities. They usually do not require assistance. They may need to make some minor modifications for certain tasks; for example, they may have to take their time climbing stairs or use a cane to get from sitting to standing position.

14. Moderate functional limitations: Patient can perform most of their daily tasks such as putting on their shoes, climbing stairs, going from sitting to standing, bathing, cooking and recreation activities with modifications. For example, they may require a higher chair with arms for sitting and a bath seat for bathing. They require assistance some of the time.

15. Severe functional limitations: Patient is unable to perform most of their daily tasks such as putting on their shoes, climbing stairs, going from sitting to standing, bathing, cooking without assistance. They require assistance most of the time. Patient can no longer participate in recreation or hobbies.

Abnormal findings on physical exam related to affected joint

16. None/mild: (Knee) Patient does not have any physical deformities such as knock knee or bowleg. (Hip) Patient is able to straighten their legs and has full range of motion. For example, their leg can be straightened when laying down. Patient can spread their legs apart.

17. Moderate: (Knee) Patient shows some level of deformity such as knock knee or bowleg. (Hip) Patient has difficulty straightening and bending their legs. For example, the patient's leg remains slightly bent when lying down. Patient has difficulty spreading their legs apart.

18. Severe: (Knee) Patient has marked deformity such as knock knee and bowleg. (Hip) Patient cannot fully bend, straighten or spread apart their legs.

Potential for progression of disease documented by radiographic findings (Primary replacement)

19. None: Patient is booked for a primary hip or knee replacement (i.e. not a revision) with no signs of protrusion or bone loss. (Note: Protrusion occurs when the ball of the joint pushes through the socket)

20. Mild: Patient is booked for a primary replacement but shows early signs of protrusion or bone loss which are barely noticeable on the X-rays.

21. Moderate: Patient is booked for a primary replacement and signs of protrusion or bone loss are clearly noticeable on the X-rays.

22. Severe: Patient is booked for a primary replacement and there are marked signs of protrusion or bone loss on the X-rays.

Threat to patient role and independence in society

23. Not threatened but more difficult: With the present level of disability, patient can continue to work with minor adjustments. With the present level of disability, patient can care for dependents but sometimes requires help. With the present level of disability, patient can continue to live independently but requires some help with jobs such as gardening and cleaning.

24. Threatened but not immediately: With the present level of disability, patient may not be able to continue to work. With the present level of disability, patient requires significant help in caring for dependents. With the present level of disability, patients living in a house with multiple floors may have to consider moving to an apartment to avoid stairs.

25. Immediately threatened or unable: With the present level of disability, patient can no longer work. With the present level of disability, patient is unable to care for dependents even with help. With the present level of disability, patient can no longer live independently and will have to be placed with relatives or moved to a care facility.

APPENDIX 3

SET 1

The table below details 8 patients. All 8 patients are **65 years old** and show **X ray abnormalities** in the left hip. All 8 have elected for primary TJR of the left hip. The patients differ in the severity of their pain and the extent of their activity limitations and social participation restrictions. You are requested to prioritise the patients for surgery. Place a **1 in the box** for the patient you consider is the **highest priority** for surgery, a 2 in the box for the patient you consider is the 2nd highest priority for surgery, and so on ending with an **8 in the box** for the patient with the **lowest priority** for surgery.

Patient 1 Severe pain Severe activity limitations Severe social participation restrictions ☐	Patient 2 Moderate pain Moderate activity limitations Severe social participation restrictions ☐
Patient 3 Severe pain Severe activity limitations Moderate social participation restrictions ☐	Patient 4 Moderate pain Severe activity limitations Severe social participation restrictions ☐
Patient 5 Severe pain Moderate activity limitations Severe social participation restrictions ☐	Patient 6 Moderate pain Moderate activity limitations Moderate social participation restrictions ☐
Patient 7 Severe pain Moderate activity limitations Moderate social participation restrictions ☐	Patient 8 Moderate pain Severe activity limitations Moderate social participation restrictions ☐

SET 2

The table below details 8 patients. All 8 patients are **65 years old** and show **X ray abnormalities** in the left hip. All 8 have elected for primary TJR of the left hip. They differ on a variety of factors detailed in the description of each patient. You are requested to prioritise the patients for surgery. Place a **1 in the box** for the patient you consider is the **highest priority** for surgery, a 2 in the box for the patient you consider is the 2nd highest priority for surgery, and so on ending with an **8 in the box** for the patient with the **lowest priority** for surgery.

Patient 1 Has severe pain Has severe activity limitations Is not motivated ☐	Patient 2 Has moderate pain Has moderate activity limitations Is highly motivated ☐
Patient 3 Has severe pain Has moderate activity limitations Is not motivated ☐	Patient 4 Has moderate pain Has severe activity limitations Is not motivated ☐
Patient 5 Has severe pain Has severe activity limitations Is highly motivated ☐	Patient 6 Has moderate pain Has severe activity limitations Is highly motivated ☐
Patient 7 Has severe pain Has moderate activity limitations Is highly motivated ☐	Patient 8 Has moderate pain Has moderate activity limitations Is not motivated ☐

APPENDIX 4 - MORNING DIARY

How many steps did you take YESTERDAY? (press MODE button until pedometer shows steps mode and then press MEMO button once to recall yesterday's data i.e. 1 day before)

How much did you walk YESTERDAY?

LESS than usual
MORE than usual

How many steps have you taken so far this morning? (press MODE button until pedometer shows steps mode)

How much have you WALKED this morning?

LESS than usual
MORE than usual

To what extent do you intend to WALK more than usual between now and the afternoon diary entry?

NO intention
DEFINITELY intend

How confident are you that you can WALK more than usual between now and the afternoon diary entry?

NOT AT ALL confident
VERY confident

How much control do you have over WALKING more than usual between now and the afternoon diary?

NO control
COMPLETE control

How much have you done GARDENING this morning?

LESS than usual
MORE than usual

To what extent do you intend to do GARDENING between now and the afternoon diary entry?

NO intention
DEFINITELY intend

How confident are you that you can do GARDENING between now and the afternoon diary entry?

NOT AT ALL confident
VERY confident

How much control do you have over doing GARDENING between now and the afternoon diary?

NO control
COMPLETE control

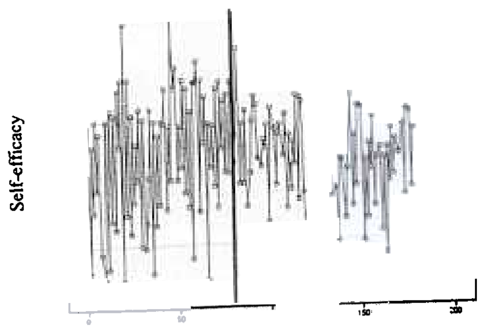
How would you describe your PAIN right now?

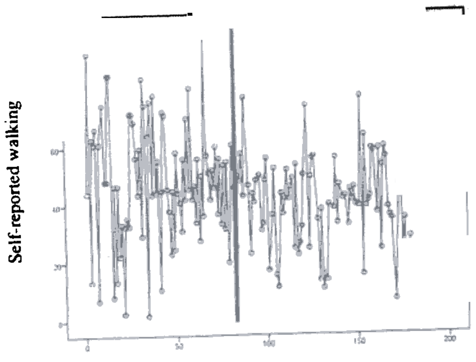
NO pain
EXTREME pain

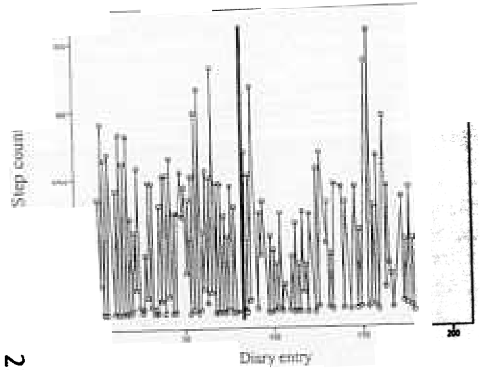
Please use this space to tell us about any important situation/event that occurred last night or this morning

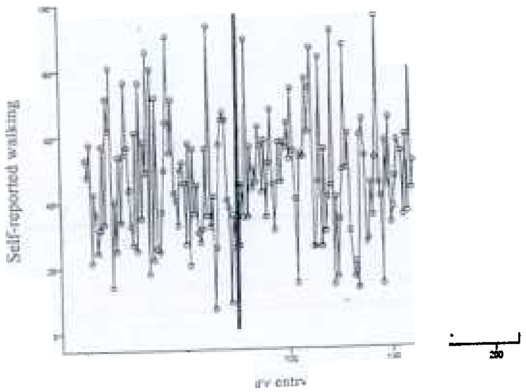
APPENDIX 5 – TIM .CTS OF IN VENTION EFFECTS

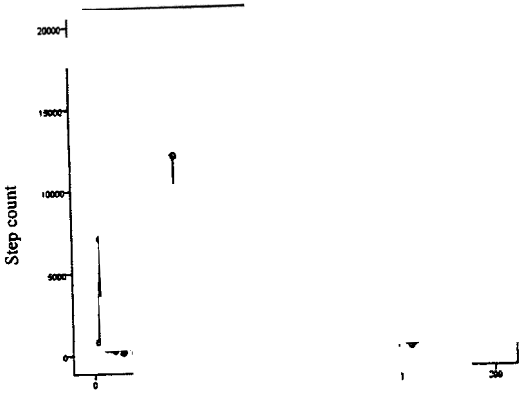
Participant 1 Self-efficacy intervention on walking





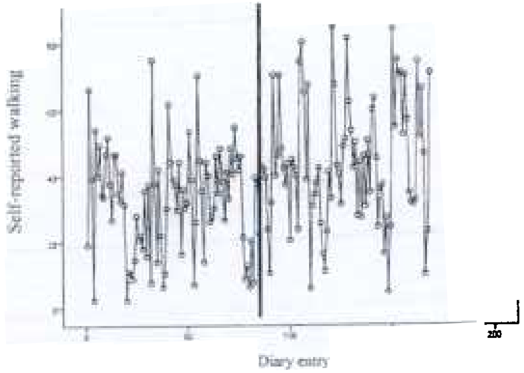


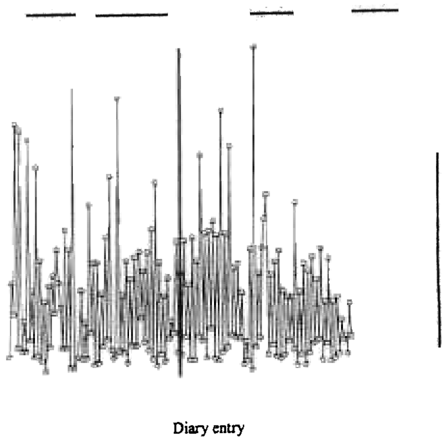




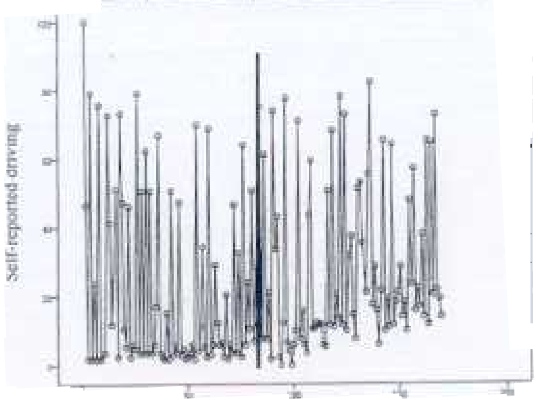
Note. Self-efficacy scale 0-100 (not at all confident - very confident); self-
physical activity behaviour scale 0-100 (less than usual - more than usual)

Participant 3 Action planning intervention on walking





Participant 4 Action planning intervention on driving



APPENDIX 5 – TEST OF INTERVENTION EFFECTS ON ALL COGNITIONS

Chosen Behaviour	Pre-intervention score[a][b]	Post-intervention score[a][b]	Statistic[a][b]
Participant 1			
SE intervention on walking			
Intention[b]	44 (83)	45 (79)	3095.5
PC[b]	97 (83)	38 (79)	2521.5*[c]
SE[a]	39 (2.0)	37 (1.3)	.99
Participant 2			
Action planning intervention on walking			
Intention[b]	32 (71)	13 (86)	2226.5**[c]
PC[b]	73 (720	76 (86)	3079.5
SE[a]	74 (1.5)	73 (1.0)	.70
Participant 3			
Action planning intervention on walking			
Intention[b]	21 (83)	20 (85)	3323.5
PC[a]	72 (0.7)	67.5 (.9)	4.15***[c]
SE[a]	64 (0.8)	62 (1.0)	1.51
Participant 4			
Action planning intervention on driving			
Intention	8 (79)	16 (87)	2485.5**
PC	87 (79)	91 (87)	1891.0***
SE	54 (79)	78 (87)	1626.5**

Note. Cognitions were measured on visual analogue scales from 0-100 (a higher score = a stronger cognition).

*p < .05. **p < .01. ***p < .001.

a pre- and post-intervention scores are mean scores (*Standard error*) and test statistic is t value.

[b] pre- and post-intervention scores are median scores (n) and test statistic is U value.

[c] Significant *decrease* in cognition score

APPENDIX 6 - DIARY

Unless otherwise stated please fill in all questions in the diary at both AM and PM entries

A: WALKING (step count)

1. To what extent do you intend to walk more than usual between now and the next time you fill in the diary?

No intention 2 3 ○ 4 6 ○ 7 8 9 10 Definitely intend to

2.
How confident are you that you can walk more than usual between now and the next time you fill in the diary?

Not at all confident 1 2 3 4 5 6 9 10 Extremely confident

3. How much control do you have over walking more than usual between now and the next time you fill in the diary?

No control 2 3 4 5 ○ 6 8 9 10 Complete control

4. To be completed at AM entry only (ignore if PM)

What is the total number of steps you took yesterday (press MODE button until pedometer shows steps mode and then press MEMO button once to recall yesterday's data i.e. 1 day before) ☐ steps

5. To be completed at PM entry only (ignore if AM)

How many steps have you taken so far today? (press MODE button until pedometer shows steps mode) ☐ steps

B: WORKING OUT IN THE GYM

1.
To what extent do you intend to work out in the gym between now and the next time you fill in the diary?

No intention 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7 ○ 8 ○ 9 ○ 10 ○ Definitely intend to

2.
How confident are you that you can work out in the gym between now and the next time you fill in the diary?

Not at all confident 1 3 4 5 8 9 10 Extremely confident

3.
How much control do you have over working out in the gym between now and the next time you fill in the diary?

No control 2 3 4 5 6 ○ 7 8 9 10 ○ Complete control

4. Have you worked out in the gym since you last filled in the diary? ○ yes ○ no

5. If you answered YES, how long did you spend working out in the gym? ☐ mins

6. If you answered YES, how intense was your workout?

No exertion at all 1 2 3 4 ○ 5 6 ○ 7 ○ 8 9 ○ 10 Extremely hard/maximal exertion

C: NON-GYM BASED PHYSICAL ACTIVITY (previously defined by YOU)

1.
To what extent do you intend to engage in your non-gym based physical activity between now and the next time you fill in the diary?

No intention ○ 2 3 ○ 4 5 6 7 ○ 8 ○ 9 10 ○ Definitely intend to

2.
How confident are you that you can engage in your non-gym based physical activity between now and the next time you fill in the diary?

Not at all confident 1 3 4 5 6 8 ○ 9 10 Extremely confident

3.
How much control do you have over engaging in your non-gym based physical activity between now and the next time you fill in the diary?

No control 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7 ○ 8 ○ 9 ○ 10 ○ Complete control

4.
Have you engaged in your non-gym based physical activity since you last filled in the diary?
○ yes ○ no

5. If you answered YES, how long did you spend engaging in this activity? ☐ mins

6. If you answered YES, how intense was your workout?

No exertion at all 1 2 3 ○ 4 5 ○ 6 ○ 7 ○ 8 9 ○ 10 Extremely hard/maximal exertion

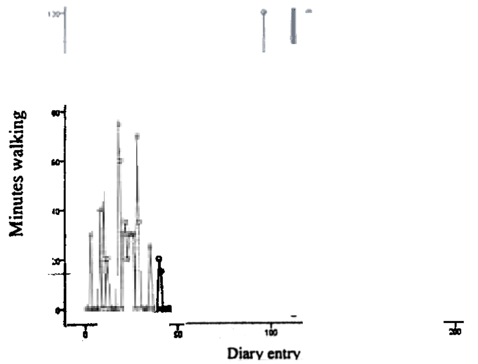
Thank you for filling in the diary.

If it is *am* now please fill it in again later today.

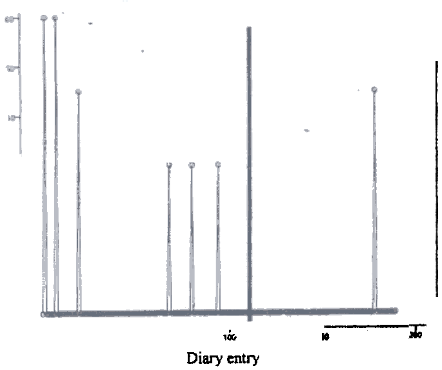
If it is *pm* now please fill it in again tomorrow morning.

APPENDIX 7 – TIME PLOTS OF IN [illegible]RVEN

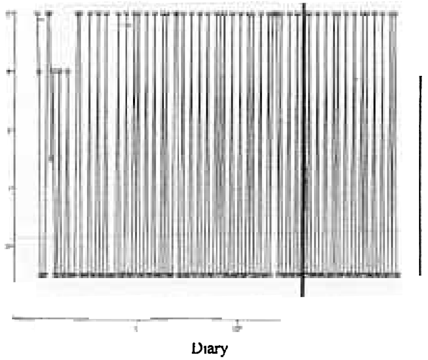
Participant 2 Action planning intervention on walk dog

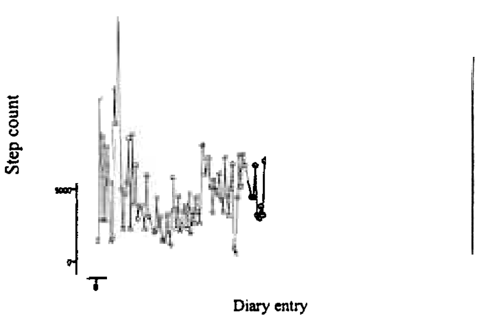


Participant [illegible] Action planning intervention on swimming

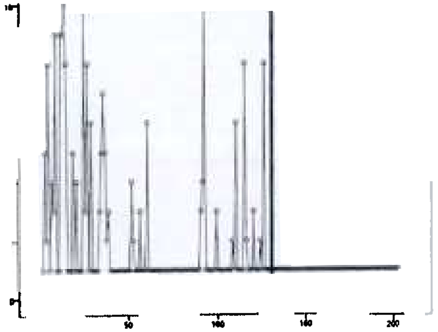


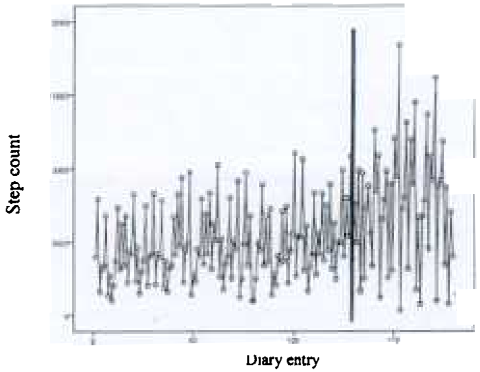
Participant 5 Self-efficacy intervention on walking





Participant 4 Perceived control intervention on gym



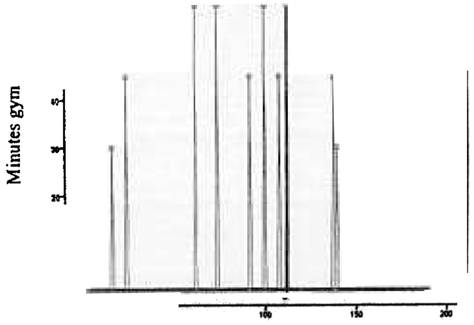


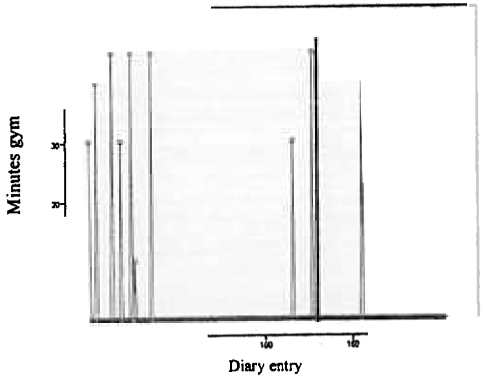
10 point scale (no control –

efficacy
all confident

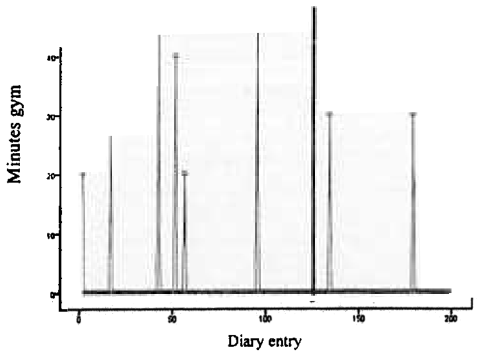
time of intervention

Participant 3 Action planning intervention on gym





Participant 6 Action planning intervention on gym



APPENDIX 7 – TEST OF INTERVENTION EFFECTS ON ALL COGNITIONS

Chosen Behaviour	Pre-intervention score[a b]	Post-intervention score[a b]	Statistic[a b]
Participant 1			
PC intervention on go for walk			
Intention[b]	5 (110)	10 (92)	3042.0***
PC[b]	5 (110)	7 (92)	1948.0***
SE[b]	5 (110)	10 (92)	2996.5***
Participant 2			
Action planning intervention on walk dog			
Intention[b]	3 (105)	1 (71)	3328.5**[c]
PC[a]	7 (0.6)	7 (0.2)	1.13
SE[b]	4 (105)	1 (71)	3266.5
Participant 3			
Action planning intervention on gym			
Intention[b]	1 (109)	1 (78)	3468.0*[c]
PC[b]	7 (108)	10 (78)	3102.0**
SE[b]	7 (108)	10 (78)	3039.0***
Action planning intervention on swimming			
Intention[b]	1 (109)	1 (78)	3286.5**[c]
PC[b]	10 (109)	10 (78)	3233.5**
SE[b]	6 (109)	7 (78)	2767.0***
Participant 4			
Perceived controllability intervention on gym			
Intention[b]	1 (131)	1 (71)	3905.0***
PC[b]	1 (131)	1 (71)	3905.0***
SE[b]	1 (131)	1 (71)	3337.0***
Participant 5			
Self-efficacy intervention on walking			
Intention[b]	1 (131)	1 (48)	3071.0
PC[b]	10 (131)	10 (48)	3057.0
SE[b]	1 (131)	1 (48)	3057.0
Participant 6			
Action planning intervention on gym			
Intention[b]	1 (115)	2 (79)	2610.0***
PC[b]	1 (116)	3 (79)	3067.0***
SE[b]	1 (116)	2 (79)	2294.5***

Note. Cognitions were measured on visual analogue scales from 0-100 (a higher score = a stronger cognition).

*p < .05. **p < .01. ***p < .001.

[a] pre- and post-intervention scores are mean scores (*Standard error*) and test statistic is *t* value.

[b] pre- and post-intervention scores are median scores (*n*) and test statistic is *U* value.

[c] Significant *decrease* in cognition score