Contributions to the Psychological Evaluation of UK Exercise Referral: The ER-QLS a Measure of Exercise-Related Life-Quality.

Charlotte Emma Hilton

A thesis submitted in partial fulfilment of the requirements of Nottingham Trent University for the degree of Doctor of Philosophy

October 2010

"Quality of life studies and measurements serve to prevent a devastating separation of a patient's body from a patient's biography during delivery of care" (Roy, 1992).

Abstract

The benefits of physical activity for health are sufficient enough to be incorporated into UK public health policy and the exercise referral scheme is a popular method of engaging sedentary individuals into a more active lifestyle. However, the evidence base as to the extent to which such schemes impact upon holistic health outcomes is limited. This deficit is particularly apparent for psychological measures, despite these being reported as the most likely outcomes of a referral into exercise. The primary reason for this problem is the lack of tools available to exercise professionals with which to capture this valuable data. This thesis proposes that, in addition to the more commonplace monitoring of physiological outcomes and physical activity level, that psychological responses to, and outcomes of, a referral into exercise should be monitored as part of routine practice. The complete thesis provides an exercise referral sensitive quality of life measure, the Exercise Referral Quality of Life Scale (ER-QLS), which is intended for use in practice and/or for the purposes of research.

Five stages of research were undertaken to produce the final measure. Two were qualitative and three were quantitative. Stage one utilised a series of five focus groups to generate rich data for the purposes of item development and a conceptual framework of exercise-related life-quality that would inform subsequent psychometric analyses. Stage two employed best-practice recommendations from previous research to construct test items and formulate a test measure that was formatted in such a way as to facilitate its completion and reduce respondent burden. Stage three cognitively pre-tested these test-pool items to ensure that they were interpreted as intended and to establish appropriate face and content validity. Stage four employed a classical testing theory approach to item reduction and also assessed the initial reliability of the measure through test-retest and internal consistency analyses. The final fifth stage employed a principal components and parallel analysis approach to exploratory factor analysis and assessed the internal consistency, test-re-test reliability, acceptability, content validity, and convergent and known groups components of construct validity of the final measure.

The ER-QLS showed good reliability and validity, is easy to administer and to score. It is expected that the measure will encourage the monitoring of exercise-related life-quality within exercise referral settings as part of more holistic approaches to evaluation. The implications for wider policy and practice are also discussed.

Acknowledgements

The process of completing this thesis has been both satisfying and challenging and there are a number of people who I would like to thank personally and professionally.

First of all, this thesis would not have been possible without the participants who kindly volunteered to help with the research at various stages. I extend my heartfelt thanks to all who gave up their time. I would also like to acknowledge the staff from the exercise referral schemes involved. This research would have been far more challenging without the co-operation and enthusiasm it received from you.

I would also like to extend my thanks to my academic mentors. In particular I would like to thank Dr Antoinette Minniti and Dr Richard Trigg for their invaluable guidance in producing the final thesis. I thank Dr Lynne Kennedy for her mentorship during the latter stages before submission and I look forward to working alongside Lynne in the future.

Special thanks go to my friend and colleague Zoe Rutherford for being such a valuable source of support through the peaks and troughs of completing this research – your turn to submit next! I would like to thank all the staff and postgraduate students at Nottingham Trent University for making my time there so enjoyable.

On a personal level I would like to extend genuine thanks to all my friends who have had to tolerate my incessant ramblings about my PhD! You're good humour and kindness have been integral to the completion of this thesis.

Thank you to my family who are a continued source of support to me. I love you all dearly and this thesis is for you.

Publications

The following publications have contained work reported in this thesis:

Conference Abstracts, Presentations and Proceedings

Hilton, C. E. &, Minniti, A. M. (2007). Exercise referral as a psychological treatment approach: Preliminary results from a qualitative investigation. *Proceedings of the American Psychological Association (APA), San Fransisco, USA.*

Hilton, C. E., & Minniti, A. M. (2007, June). *Towards a standardised toolkit for exercise referral. Introducing quality of life assessment.* Paper presented at the Nottingham Trent University Research Conference, Nottingham, UK.

Hilton, C. E., & Minniti, A. M. (2008). Responses to and experiences of exercise referral: A focus group approach [Abstract]. *Proceedings of the European College of Sport Sciences, Estoril, Portugal.*

TABLE OF CONTENTS

Abstract	**I**
Acknowledgements	**II**
Publications	**III**
TABLE OF CONTENTS	**IV**
List of Tables	**IX**
List of Figures	**XI**
CHAPTER I Background and Review of the Literature	**1**
1.1 Purpose of the Thesis	2
1.2. General Introduction	4
1.3 Exercise Referral Systems: A National Quality Assurance Framework	6
1.3.1 Controlled Studies	8
1.3.2 Audits	8
1.3.3 Reflective Practice	9
1.4 Exercise Referral Schemes and the Evaluation Chameleon	11
1.5 Moving Towards the Holistic Assessment of Exercise Referral	26
1.6 The Conceptual and Assessment Challenges of Life-Quality	29
1.7 The Exercise and Quality of Life Relationship	35
1.8 Health as a Holistic Construct: Matching the Tools to the Task	44
1.9 Summary of the Literature Review	47
CHAPTER II Scale Development: Overview and Focus Groups	**49**
2.1 Scale Development: Overview	50
2.1.1 Introduction: Scale Development – Focus Groups	52
2.1.2 The Focus Group Approach	54
2.1.3 Data Collection	54
2.1.4 Reflexivity	57
2.1.5 Approaches to Data Analysis and Representation	57
2.1.6 Rigorous Qualitative Research	58
2.2 Aims	60
2.3 Methods	60
2.3.1 Participants	65
2.3.2 Data Handling and Analysis	68

2.4 Results: Open Coding 75
2.4.1 Results: Higher Order Domains 98
2.4.2 Reflections of the Interviewer 104
2.5 Discussion 105
2.5.1 Limitations of the Research 111
2.6 Conclusions 112
CHAPTER III Scale Construction: Phases One and Two 113
3.1 Introduction: Item and Questionnaire Design 114
3.2 Aims 117
3.3 Methods 117
3.4 Resultant Items and Scales 119
3.5 Introduction: Cognitive Pre-Testing 124
3.6 Aims 126
3.7 Methods 126
3.7.1 Data Handling and Analysis 128
3.7.2 Participants 129
3.8 Results 131
3.9 Reflections of the Interviewer 136
3.10 Discussion 138
3.10.1 Limitations of the Research 142
3.11 Conclusions 142
CHAPTER IV Phase Three: Item Analyses, Reduction & Initial Reliability 144
4.1 Introduction: Item Analyses, Reduction and Reliability 145
4.2 Aims 149
4.2.1 Objectives 149
4.3 Methods 150
4.3.1 Participants 150
4.3.2 Materials 154
4.3.3 Procedure 161
4.3.4 Analyses 161
4.4 Results 163
4.5 Discussion 179
4.5.1 Limitations of the Research 183
4.6 Conclusions 183
CHAPTER V: Phase Four: The Factor Structure, Reliability, Validity 185
and Acceptability of the Final Scale 185

5.1 Introduction: Factor Analysis 186
5.1.1 Exploratory Factor Analysis 188
5.1.2 Confirmatory Factor Analysis 188
5.1.3 Principal Components Analysis 189
5.1.4 Factor Rotation 189
5.1.5 Reliability, Validity and Acceptability 190
5.2 Aims 192
5.3 Methods 192
5.3.1 Participants 193
5.3.2 Materials 194
5.3.3 Selection of Construct Measures and Correlation Hypotheses 196
5.3.4 Procedure 199
5.3.5 Analyses 200
5.4 Results 201
5.4.1 Principal Components Analysis 201
5.4.2 Parallel Analysis 204
5.4.3 Content Validity 205
5.4.4 Reliability 210
5.4.5 Construct Validity 212
5.4.6 Acceptability 215
5.4.7 Organization of Items in the Final Draft Scale 216
5.5 Discussion 219
5.5.1 Limitations of the Research 225
5.6 Conclusions 225
CHAPTER VI Properties of the Final Scale and Scoring Protocol 227
6.1 Introduction 228
6.2 Appropriateness and Feasibility 228
6.2.1 Acceptability 230
6.2.2 Precision 230
6.2.3 Reliability 231
6.2.4 Face and content validity 232
6.2.5 Construct Validity 233
6.2.6 Reproducibility 234
6.2.7 Interpretability 234
6.3 Appropriate Scale Title: The ER-QLS 235
6.4 Scoring the ER-QLS 236

6.5 Instructions for Administrators 239
6.6 Discussion 240
6.6.1 Item Generation 241
6.6.2 Item Construction 242
6.6.3 Cognitive Pre-testing 242
6.6.4 Item and Scale Analyses 244
6.7 Conclusions 246
CHAPTER VII Discussion and Conclusions 247
7.1 Introduction 248
7.2 The ER-QLS and the Measurement of Life Quality 248
7.3 Exercise-Related Life-Quality 250
7.4 Beyond Cause and Effect 256
7.5 Strengths and Limitations of the Research Programme 259
7.6 Future Research Directions 262
7.6.1 Qualitative Research and Theoretical Contributions 262
7.6.2 Further Validation of the ER-QLS and Longitudinal Study 264
7.7 Implications for Policy and Practice 266
7.8 Conclusions 268
References 270
Appendices 305
Appendix 1 NRES Participant Information Sheet: Focus Groups 306
Appendix 2 Information for Participants : Focus Group 312
Appendix 3 Consent Form 313
Appendix 4 NVivo Node Report 314
Appendix 5 NRES Participant Information Sheet: Questionnaire Pre-Testing 317
Appendix 6 Pre-Testing Information for Participants: Self Complete 321
Appendix 7 Pre-Testing Information for Participants : Interview/Telephone 322
Appendix 8 Interviewer Probe Protocol 1 323
Appendix 9 Interviewer Probe Protocol 2 329
Appendix 10 Interviewer Probe Protocol 3 336
Appendix 11 Pre-Testing Outcome Report: Self Complete 343
Appendix 12 Pre-Testing Outcome Report: Interview-Administered 349
Appendix 13 Pre-Testing Outcome Report: Telephone Administered 355
Appendix 14 Test Questionnaire 361
Appendix 15 NRES Participant Information Sheet: Questionnaire Validation 369
Appendix 16 Questionnaire Validation Information for Participants 373

Appendix 17 Cover Letter to Recruitment Sites 374
Appendix 18 Administration Protocol for Recruitment Sites 375
Appendix 19 The Final Scale: The ER-QLS 376
Appendix 20 SPSS Syntax File for Scoring the ER-QLS 380

List of Tables

Table 1. Levels of evaluation recognised by the National Quality Assurance Framework (DH, 2002). 7
Table 2. Demographic characteristics of exercise referral focus group participants (*N=23*). 67
Table 3. Open coded data and the phase of exercise referral at which it was reported 76
Table 4. Coded categories grouped into the domain of leisure facilities and lifestyle physical activity 99
Table 5. Coded categories grouped into the domain of confidence and motivation 100
Table 6. Coded categories grouped into the domain of discomfort and fitness 100
Table 7. Coded categories grouped into the domain of mental and physical well-being 101
Table 8. Data not included in the higher order analysis and retained separately 103
Table 9. Questions reflecting the construct leisure facilities and lifestyle physical .120
Table 10. Questions reflecting the construct confidence and motivation.....121
Table 11. Questions reflecting the construct Discomfort and fitness 121
Table 12. Questions Reflecting the Construct: Mental and Physical Well-being 122
Table 13. Pre-testing participants demographic data.....130
Table 14. Summary of pre-testing results: Self-complete, interview and telephone administered.....132
Table 15. Number of respondents recruited from each test site 151
Table 16. Wording of items representing the construct of leisure facilities and lifestyle physical activity 157
Table 17. Wording of items representing the construct of confidence and motivation 158
Table 18. Wording of items representing the construct of discomfort and fitness...159
Table 19. Wording of items representing the construct of mental and physical well-being 160
Table 20. Summary of item analysis of the test-pool (N = 50) 173
Table 21. Retained items, domains and revised item-total correlations (N = 29) 177
Table 22. Items included in the principal components analysis (N = 29) 195
Table 23. Final 22-item three-factor solution and domain titles 203
Table 24. Eigen and parallel analysis values of the three-factor solution (N = 22) .205
Table 25. Items rejected from the original 50-item test pool 207

Table 26. The factor and conceptual framework component of final items.....208
Table 27. Factor loadings of a single factor solution (N = 22).....209
Table 28. Item-total correlations and Cronbach's alpha values of the 22-item three-factor solution.....211
Table 29. Descriptive statistics of the test measure total score and domain scores of the SEES and WHOQOL-BREF.....212
Table 30. Descriptive statistics for total dimension scores of the test measure WHOQOL-BREF and SEES.....213
Table 31. Sub-domain correlations between the test measure, WHOQOL-BREF and SEES.....214
Table 32. Descriptive statistics for known group validity analyses.....215
Table 33. Question order of final items.....217
Table 34. Percentile scores for the ER-QLS.....239

List of Figures

Figure 1. Procedures undertaken to develop and validate the final scale 51
Figure 2. Causal and effect indicators (Streiner & Norman 2008). 71
Figure 3. Screen print of the description of the affordability NVivo node.....72
Figure 4. NVivo screen print of the memo indicating the psychological significance of better health. 73
Figure 5. NVivo screen print of the memo indicating potential links between coded data and theoretical frameworks 73
Figure 6. Screen print example of an NVivo electronic diary entry: Potential research questions 74
Figure 7. Screen print example of an NVivo electronic diary entry: Development of the coded data 74
Figure 8. A conceptual framework of exercise-related life-quality.....102
Figure 9. Example of the response scale to assess frequency.....123
Figure 10. Example of the response scale to assess capacity 123
Figure 11. Example of the response scale to assess evaluation perceptions 123
Figure 12. Number of items representing the four conceptual domains of exercise-related life-quality.....156
Figure 13. Scree plot indicating a three-factor solution 202
Figure 14. A three factor representation of exercise-related life-quality.....218
Figure 15. Measurable indicators of exercise-related life-quality 255

CHAPTER I Background and Review of the Literature

1.1 Purpose of the Thesis

The current series of studies were undertaken with a primary aim to connect practitioners based within exercise referral settings with evidence-based tools and insights grounded in scientific research. The thesis adopts a health psychology disciplinary perspective that is closely related to matters of concern for public health. It was considered that a health psychology perspective would be most suited to the current research because the primary purpose of the thesis was to respond to the psychological assessment and evaluation shortfall currently observed within exercise referral settings in the UK. The following section of the thesis includes: an overview of the purpose of the research followed by a summary of current issues relevant to the purpose; and finally a comprehensive review of the literature that both encompasses a critique and analysis of related government policies/strategies and is balanced in light of academic research that has significance for exercise referral scheme delivery and evaluation.

The purpose of the current studies is to respond to the current challenges that are met by exercise professionals working within adult exercise referral settings. The data collection tools that are available with which to monitor the holistic health outcomes of referral into a 12-week exercise programme are poorly suited to the needs of exercise professionals. Practitioners are often reliant upon physiological parameters that, while important, cause difficulties in terms of evaluating scheme impact from a more complete health perspective. Significant changes in physiological parameters may take longer than an initial 12-week referral into exercise, particularly if those referred choose to exercise once to twice weekly during the course of their programme rather than three times or more. Yet, physiological outcomes are the most commonly monitored responses to and outcomes of exercise referral by exercise professionals working in the field. Attempts to monitor psychological outcomes, if attempted at all, usually take the form of generic measures of quality life (QoL). The most commonly used instruments are the Short Form 36 (SF-36; Brazier et al., 1992; Ware, 1996) and the EuroQoL (EQ-5D; Brooks, 1996). However, practitioners often report these measures as being ill fitted to their needs in terms of length of the questionnaire, appropriateness and relevance of questions and challenging scoring protocols. In these instances it is commonplace for practitioners to utilise a population specific measure that usually reduces or eliminates some of

these difficulties. These are often referred to as health-related measures of life-quality. However, such a measure for exercise referral is currently unavailable.

Therefore, the primary research undertaken for this thesis was for the development of a quality of life (QoL) measure. The measure has been designed to be sensitive to those areas of QoL that are of particular importance and relevance to those individuals undertaking a programme of exercise for health improvement. The development of the measure is intended to contribute to remedying the challenges faced by exercise professionals whose objective is to evaluate more holistically, beyond and in addition to physiological parameters.

The QoL assessment measure is not intended to be prescriptive. Rather, the intention is to contribute to the growing need to evaluate exercise referral schemes more effectively (e.g., NICE, 2006a) and to provide practitioners with the most appropriate tools for the task. Equally, it is expected that the provision of a measure of exercise-related life-quality will support the notion of and scope for some level of standardisation and consistency for exercise referral programme design, delivery and evaluation and, in turn, this will lead to improvements in the clarity of data that intends to demonstrate the impact of schemes upon holistic health.

Throughout the thesis, the terms physical activity (PA) and exercise are used. PA is recognised as a general term that refers to any movement of the body that results in energy expenditure above that of resting level (Caspersen, Powell, & Christenson, 1985). Within the literature, the terms PA and exercise are often used interchangeably. However, it is acknowledged that exercise refers to a component of PA that is more structured and purposefully undertaken for example, gym attendance, jogging, or attending an exercise class (e.g., Caspersen et al., 1985).

1.2. General Introduction

The benefits of physical activity (PA) for health and well-being are well documented (e.g., Biddle & Mutrie, 2001; Brukner & Brown, 2005; Warburton, Nicol, & Bredin, 2006) and the consequences of physical inactivity and sedentary living have also been addressed from both a health implication and an economic cost perspective (e.g., Department of Health (DH), 2004b; Garrett, Brasure, Schmitz, Schultz, & Huber, 2004). It is estimated that approximately 60% of men and 70% of women in the UK are not sufficiently active enough to benefit their health (DH, 2004b). As a result, PA is increasingly being considered the best investment in health and has been included into recent public health policy (Bull & Milton, 2010; DH, 2004a; 2004b; 2005; 2009a; National Institute for Health & Clinical Excellence, 2006a; 2006b; 2006c).

A UK report from the Chief Medical Officer (DH, 2004b) communicated that adults should be accumulating 30 minutes of at least moderate intensity physical activity on five or more days a week in order to maintain good general health (5 x 30 moderate). This amount is also reported as sufficient enough to improve psychological well-being although 45-60 minutes of moderate intensity physical activity everyday is recommended for the prevention of obesity. The report acknowledges the significant role that physical activity has in the prevention and management of up to 20 chronic diseases and the resultant financial burden that physical inactivity has on the economy. In England, the cost of physical inactivity to the economy is estimated to be around £8.2 billion with an additional cost of £2.5 billion for the cost of obesity alone (DH, 2004b). The role of physical activity in terms of disease prevention, management, improved mental well-being and the potential economic savings results in physical inactivity being considered as a major public health issue.

Health Survey for England data show that in 2003 only 37% of men and 24% of women met the current physical activity guidelines suggested by the Government (DH, 2005). More recent Health Survey for England data show that 40% of men and 28% of women met the current physical activity guidelines (DH, 2006) and in 2008 these levels rose marginally to 46% of men and 36% of women meeting the guidelines (DH, 2008). In 2002, the Government proposed to increase the proportion of the English adult population who participated in 30 minutes of moderate physical

activity five or more times a week to 70% by 2020. This would require participation levels in England to double in just over 15 years. As a result, physical activity is increasingly considered as the best investment in health and has been included into recent UK public health policy (DH, 2004a; DH, 2004b; DH, 2005; DH, 2009a; National Institute for Health & Clinical Excellence, NICE; 2006a; 2006b; 2006c).

The National Institute for Health and Clinical Excellence (NICE) offered a review of four commonly used methods to increase PA, namely: a) brief interventions in primary care, b) exercise referral schemes, c) pedometers and d) community based programmes for walking and cycling (NICE, 2006a). Exercise referral schemes are typically designed to offer a programme of exercise for around 12 weeks, and in the UK guidelines have been published to assist organisations with scheme design and evaluation (Exercise Referral Systems: A National Quality Assurance Framework [NQAF] (DH, 2002). The recommendation for exercise referral schemes was that they should only be endorsed when part of a properly designed and controlled research study to determine effectiveness (NICE, 2006a) and an adjacent review of exercise referral schemes concluded that they were ineffective in increasing PA in the long-term (NICE, 2006c).

Indeed, there has been much debate regarding the efficacy of exercise referral schemes in terms of their ability to increase individual PA levels, and their contribution to health improvements (e.g., Dugdill, Graham, & McNair, 2005; Harrison, Roberts, & Elton, 2004; Morgan; 2004 Riddoch, Puig-Ribera & Cooper, 1998). However, significantly more attention is given to the changes in PA levels of exercise referral participants and physiological outcome measures, despite psychological measures being "more likely to change than physical measures over a 10-week exercise programme" (DH, 2002, p. 44). Generally, relevant psychological and environmental parameters tend to be ignored (Dugdill et al., 2005) and the evidence base to support the efficacy of exercise referral schemes is often regarded as conflicting or inconclusive (Crone, Johnston, & Grant, 2004; Graham, Dugdill, & Cable, 2005). However, as with the NICE (2006c) review, the literature tends to demonstrate a narrow focus on participant outcomes and usually employs a randomised control trial (RCT) methods design (e.g., Isaacs et al., 2007; Lawton, Rose, Elley, Dowell, Fenton, & Moyes, 2008; Taylor, Doust, & Webborn, 1998).

Despite these inconsistencies, the exercise referral scheme remains one of the most prolific initiatives developed (Crone et al., 2004). With the growing number of PA and exercise initiatives that aim to tackle public health issues, there is a critical opportunity to conduct research that serves to facilitate the quality of exercise referral schemes (McKay, Macdonald, Reed, & Khan, 2005). There is considerable political emphasis placed on the role of PA for health (Crone-Grant, 2001), and there has never been a more fruitful arena in which to conduct research that aims to identify the diversity of reported responses to and outcomes of PA and exercise for those who have been signposted to structured and supervised opportunities by health professionals.

The objective of this thesis is to improve the tools available to exercise professionals with the aim of supporting consistent yet holistic delivery and evaluation of exercise referral schemes. An influential UK document published by NICE (2006a) identified four commonly used methods to increase PA. The document sought to consolidate the evidence-base for the efficacy of these four methods to increase PA. However, given the focus of the current thesis is to contribute to the holistic evaluation of UK exercise referral schemes, the following review will focus upon the exercise referral component of the NICE (2006a) publication and exercise referral literature specifically. The UK published guidance for professionals regarding the best practice design and evaluation of schemes will be presented (DH, 2002) and the inconsistent nature of exercise referral scheme evaluation will be critiqued with references to recent research and the relevance for applied practice. The implications of integrating a holistic evaluation framework into exercise referral settings while recognising the challenges of QoL measurement and the strength of the exercise and QoL relationship will be discussed. The importance of matching appropriate evaluation tools to the task of evaluating diverse health outcomes will be explored within the context of current understandings regarding the notion of health.

1.3 Exercise Referral Systems: A National Quality Assurance Framework

During the 1990s, the number of exercise referral schemes in the UK developed rapidly (Fox, Biddle, Edmunds, Bowler, & Killoran, 1997; Graham, Dugdill, & Cable, 2005) and as a consequence, it was difficult for scheme developers, co-ordinators and stakeholders to determine a best practice model, or at least a minimum

standard for exercise referral protocol design, delivery and evaluation. The UK government response was to develop best practice design and evaluation guidelines in an attempt to introduce a level of continuity and professional standards.

The National Quality Assurance Framework (NQAF) (DH, 2002) provides guidelines for exercise referral schemes with the intention of improving standards for existing schemes and those under development. The scope of the document covers a range of considerations that are important at each stage of the referral pathway process. For example, areas addressed include the screening and selection process for those referred, choice of activities, the competencies of exercise professionals delivering the programme, evaluation mechanisms and long-term support.

The NQAF (DH, 2002) recognises three levels of evaluation, namely: a) controlled studies, b) audits and c) reflective practice (Table 1). The document provides recommendations and limitations of each evaluation level and are intended to act as guidelines for exercise referral professionals.

Table 1. Levels of evaluation recognised by the National Quality Assurance Framework (DH, 2002).

Level	Personnel Involved	Stakeholder	Main Function
Controlled study	Experienced researcher	Health service commissioners and policy makers	To inform health analysts, policy makers and managers
Audit	Trained referrer and/or health professional commissioners	Health service commissioners or leisure providers	To inform health service commissioners or leisure providers
Reflective practice	Referrer and/or exercise professional	Referrer and/or exercise professional	To develop a reflective (critical) professional

1.3.1 Controlled Studies

The randomised controlled trial (RCT) is often proposed as the gold standard of a controlled study. However, it is recognised that this method may not be feasible for exercise referral schemes because the least active individuals are unlikely to enrol in a study or complete follow-up assessments, and there are difficulties with randomly assigning patients (Dugdill et al., 2005). Better adherence rates may be observed in smaller scale schemes because those who are referred usually benefit from support from specific exercise professionals (Taylor, 1998). Similarly, the NQAF document recognises the intense resource implications of the RCT method in terms of researcher training, equipment and sophisticated data analysis procedures.

A single evaluation guideline for controlled trials is offered, stipulating that resources should be allocated to conduct trials with clear objectives and that research expertise in health and exercise science, health and exercise promotion and health economics should be utilised. Perhaps one of the main limitations of an RCT approach to evaluation, particularly within community settings is the ethical consideration of who is selected to participate in the experimental condition (in this case an exercise referral scheme) and potentially benefit their health, and who is not. In this respect, such an approach to scheme evaluation poses sensitive ethical considerations and is not congruent with tackling inequalities in health. Dugdill et al. (2005) have argued that the recommendation of an RCT approach should be viewed as one of the main criticisms of the NQAF (DH, 2002) document and instead call for a greater exploration of the types of research methodologies that could be used to evaluate exercise referral schemes more effectively.

1.3.2 Audits

The NQAF audit recommends long-term follow up assessments of physical activity change but with the stipulation that the accuracy of such assessment requires an understanding of type, frequency, intensity and duration of exercise. Process measures such as the perceptions of the referring professional, patient and other exercise facility users are advised. The use of qualitative methods to inform a complete evaluation are recommended – such as individual patient interviews or focus groups. However, the document recognises that training in the use of process,

outcome and qualitative measures may be necessary. Suggestions for outcome measures that may be audited include physical, behavioural, social and psychological measures. Indeed, practitioners are advised that "psychological measures may be more likely to change than physical measures over a 10-week exercise programme" (DH, 2002 p.44).

Six evaluation guidelines for audits are offered, including: 1) all exercise referral systems should have an integral auditing system designed around agreed outcomes between health and exercise professionals and commissioners with a clear mechanism for information exchange; 2) audited measures should consider patient satisfaction based around physical factors, lifestyle factors, medication use and other treatments and psychological and social outcomes; 3) training in process data collection should be provided and incorporated into a service level agreement to ensure support and continuing professional development for the practitioner; 4) audited measures should be patient-centred, that is, understood by patients and used as part of a motivational strategy that includes goal setting, feedback and enhancement of patient perceived competence (self-efficacy); 5) audited measures are easily obtainable and include long-term change in PA (nine months and beyond), and these should be included into the normal operation cost of a scheme; and 6) data should be pooled to allow for analysis of the specific determinants of adherence to exercise referral programmes and long-term behaviour change.

While these guidelines are perhaps intended to broaden the manner in which the evaluation of schemes is conducted, in practice addressing these guidelines may not be feasible. Conducting an evaluation by employing these suggested audit strategies requires quite a varied skill-set that typically goes beyond those of exercise referral personnel (Dugdill et al., 2005). Moreover, the notion of including motivational strategies and enhancement of self-efficacy sit very much within the psychological measurement domain that (as already mentioned) are particularly lacking within current evaluation practice.

1.3.3 Reflective Practice

Collaborative action research is suggested as a method of modifying practice, changing practice and reflecting upon the changes made. The Register of Exercise

Professionals (REPS) is an organisation developed with the purpose of identifying exercise practitioners' qualifications and their skill sets with a view to classifying their level of expertise. In order to continue with registration, exercise professionals must demonstrate continuing professional development and the NQAF suggests that this requires an element of reflective practice.

In accordance with NQAF guidelines, it is critical that instructors embrace psychological determinants of exercise participation and outcomes. The mechanisms by which the NQAF recommendations may be implemented by exercise and health professionals must be given more attention in the literature, as this information is currently lacking (Crone, Johnston, & Grant, 2004). Practical recommendations for schemes and professionals as to how the most appropriate skills may be developed, implemented, maintained and improved need to be provided. Finally, appropriate and validated tools need to be developed to enable exercise professionals to collect meaningful data that is easily communicated between multi-agency health and leisure groups.

Two evaluation guidelines for reflective practice are offered, including that: 1) continuing professional development is encouraged to facilitate self-reflection and 2) trained professionals offer guidance regarding the effectiveness of personal practice in terms of interpersonal and communication skills.

The NQAF (DH, 2002) is central to the design and evaluation of UK exercise referral schemes and went some way to clarifying what was expected in this respect. Exercise referral schemes across the UK have responded well however, the evaluation of scheme efficacy in terms of outcomes for participants, in particular, remains challenging. Perhaps the most appropriate research designs for the evaluation of exercise referral schemes are ones that are able to honour the more complete understanding of health that includes the consideration of an individual's biology, psychology and social circumstances (Engel, 1977). In order to achieve this, a range of research skills and tools would be needed that have the ability to both qualitatively and quantitatively assess not only scheme impact (outcome data) but more process related factors also. Additionally, such research methods would be required to honour a commitment to ethical considerations and reducing inequalities in health which is one of the main criticisms of the RCT approach within this

context. It has been suggested that increasing PA, reducing health inequalities and enhancing QoL requires a multi-disciplinary approach (Hilton, 2010). In this respect the evaluation methods used to assess how well exercise referral schemes have contributed to these aims should also be varied in nature, yet fit for purpose.

It has been noted previously that the success of schemes is often assessed in view of the number of individuals referred and that this approach limits issues of quality and sustainability (Dugdill et al., 2005). Variations in scheme funding streams, resources, expectations of stakeholders, personnel skill-set and training and the availability of measurement tools in particular, often result in very different evaluation protocols from one scheme to another. Furthermore, any data that are collected can often be of limited use because the measurements that are taken are inappropriate in terms of informing scheme improvement (Dugdill et al., 2005). While it is important to evaluate interventions within the scope of the specific aims of that intervention and the resources available locally, practitioners are still left wanting in terms of monitoring and reporting the psychological outcomes for those referred.

1.4 Exercise Referral Schemes and the Evaluation Chameleon

It is important to critically explore the typical design, delivery and evaluation of exercise referral schemes in the UK before reviewing the literature that relates to constructs of health and the life-quality-exercise relationship. Doing so will ensure that the rationale for introducing an exercise referral sensitive QoL measure into such settings is more apparent. Furthermore, the literature within the domain of health constructs and exercise related life-quality will have greater significance when presented within the context of current exercise referral practice. Therefore, what follows is a review of the similarities and differences of scheme designs and a critical exploration of the methods and tools used to evaluate patient health outcomes. The notion of quality control issues across schemes will also be assessed in light of relevant published research.

Despite the guidance provided by the NQAF (2002), the design, and more commonly the evaluation of UK exercise referral schemes can differ greatly. However, there are commonalities. The design of a typical UK exercise referral scheme includes a ‘patient pathway’ that allows health professionals (usually General Practitioners and

Practice Nurses) working within primary care to refer patients for programmes of structured exercise. Each scheme will have a specific set of medical inclusion criteria, usually based around particular levels of Body Mass Index (BMI) or Blood Pressure (BP), for example. The referring health practitioner has the ability to signpost eligible patients to a scheme whereby it is commonplace for a scheme coordinator to identify appropriate opportunities for physical activity based upon medical history, disease profile and patient choice. For those who are signposted to structured exercise within a gym setting, instructors supporting these individuals should hold a qualification to work with special populations (e.g., WRIGHT Foundation Exercise Referral).

Exercise referral schemes are typically designed to offer a programme of exercise for around 12 weeks, and in the UK guidelines have been published to assist organisations with scheme design and evaluation (DH, 2002). However, despite these guidelines, it is the comprehensive evaluation of schemes that proves the most challenging in practice and generates inconsistencies in evaluation methods alongside a physiologically weighted focus upon scheme impact. It has also been recognised previously that the evaluation guidelines provided by the NQAF are limited which may account for practitioners feeling unsure of the best measures to use when assessing effectiveness (Dugdill et al., 2005).

The perceived difficulties and barriers to quality evaluation practices have been described in terms of cost implications, time restrictions, a method of delaying decisions or justifying cutbacks and even something that is just too difficult (Springett & Dugdill, 1999). In practice, the nature of scheme evaluation is often as wide ranging as the protocol for the scheme itself. Factors that are deemed as appropriate measures of success are often influenced by the interests of organisations who fund the respective schemes. These factors can often be very different and, in the main, future funding is usually dependent upon satisfactory evidence of participant (physical) health improvement and rates of referral. In addition, the specific skills and experience of scheme personnel and scheme resources are often guiding factors that influence the type of scheme data is collected for the purposes of monitoring and evaluation, if any. In the context of exercise referral this data is usually physiologically orientated.

Traditionally, healthcare has been deeply rooted within a medical model of practice which may account for the minimal consideration of psycho-social experiences and outcomes of those referred for PA and structured exercise. Gidlow and Murphy (2009) have noted that studies that have not adopted a RCT approach to scheme evaluation tend to be excluded from reviews that inform best practice. A further consideration is that funding and facilitation of such schemes is often provided by a local Primary Care Trust (PCT). Perhaps as a consequence, there has been a greater emphasis on the value of clinical, quantitative data collection, despite calls for broader, more holistic evaluations of exercise referral schemes (Dugdill et al., 2005; Gidlow & Murphy, 2009; Oakley, 2001; Springett & Dugdill, 1995; Taylor, 2003).

Recommendations to move away from the 'physiological measurement only' model, and encourage quantitative *and* qualitative methodology to enable more meaningful measurement of quality experience and health outcome from all stakeholder perspectives have been made (e.g., Dugdill et al., 2005). More importantly, these types of studies can recognise the value of person-centred methods of healthcare practice, and focus on the person with the disease rather than the course of the disease itself (Bauman, Fardy, & Harris, 2003). Furthermore, quality of life (QoL) measures have been recognised as a valuable method of addressing a shift in attention from controlling a patient's disease to their experience of potential suffering (Cohen, Mount, & MacDonald, 1996).

In terms of increasing physical activity levels, the efficacy of exercise referral schemes particularly in the longer-term is poor (NICE, 2006c). Although this is largely due to a lack of available data post 12 months. Only four randomised controlled trials of exercise referral schemes were included in a review by NICE (2006c), it was deemed that such schemes should only be considered beneficial in increasing physical activity levels in the short-term (6-12 weeks) and that in the longer term (over 12 weeks) and very long-term (over one year) schemes are ineffective. However, the NICE (2006c) data highlights one of the limitations of the experimental data available regarding the evaluation of exercise referral schemes. Specifically, that such data only provides insights into increased PA and not how any such increases may have been achieved (Gidlow & Murphy, 2009).

Perhaps this highlights the limitation of exercise referral schemes to support long-term behaviour change which is where the strengths of brief interventions lie and as

such have been supported by NICE (2006a). However, it is important to be mindful that current measures of physical activity are not designed to detect change over time which is problematic when attempting to quantify this measure without the use of tools such as accelerometers for example, which may indeed present their own measurement difficulty or when those responsible for assessing patient outcomes are reliant upon self-reported physical activity level.

In practice, those responsible for the evaluation of exercise referral schemes do not have access to objective yet expensive measures of physical activity - for example, global positioning system (GPS) or accelerometer equipment. Pedometers are sometimes introduced as a cheaper alternative although the accuracy of measurement has been questioned, particularly in terms of comparisons to the more expensive accelerometer (Le Masurier & Tudor-Locke, 2003). Moreover, the evaluation guidelines provided by the NQAF (DH, 2002) are limited (Dugdill, et al., 2005) and as a consequence, practitioners often struggle to select the most appropriate outcome measures to detect the diverse outcomes of a referral into exercise. Subsequently, an increasing number of schemes seek to collaborate with local universities in an attempt to improve evaluation procedures and the quality of data collected. This may also contribute to fulfilling the evaluation recommendations communicated by NICE that stipulate an exercise referral scheme should only be endorsed by practitioners, policy makers and commissioners when part of a properly designed and controlled research study to determine effectiveness (NICE, 2006a).

What follows is a review of previous studies that have evaluated exercise referral schemes. Given that the focus of the current thesis is to support the evaluation of UK exercise referral schemes, the inclusion criteria for studies was that the research must have been conducted on a scheme which was typical of UK exercise referral schemes outlined by the NQAF (DH, 2002). Although there is variance in the methods used to assess scheme performance and patient outcomes, for this section of the review there were also specific exclusion criteria; i.e., studies that included resources or support mechanisms for participants/personnel that went beyond those usually at the disposal of exercise referral schemes in the UK (for example, added resource support from the research team or specialist training for scheme personnel) were not reviewed here because such studies would not be representative of the schemes that the current research aims to support. Furthermore, with respect to specialist training for

personnel, although the long-term aim of exercise referral schemes is to encourage sustainable behaviour change, the inclusion of behaviour change strategies into scheme delivery is often (but not exclusively) reserved for experimental trials rather than part of routine practice. Schemes that have been specifically designed to address behaviour change are reviewed later.

One of the earliest and commonly cited investigations into the effectiveness of primary care PA promotion schemes was conducted by Riddoch et al. (1998), the findings of which offer direct support for the need for a psychological and person-centred approach to scheme delivery and evaluation. Riddoch et al. (1998) utilised a systematic review approach but also incorporated three case studies of existing schemes into the study design. Although exercise referral schemes were not the sole focus of the investigation, the inclusion criteria for the studies reviewed by Riddoch et al. (1998) reflect those design components that typify UK exercise referral schemes (i.e., eligibility of adults > 16 years, primary aims of the scheme being to improve PA levels, mediators of PA or attitudes/intentions towards activity and that the scheme was initiated in primary care). Consequently, the findings of Riddoch et al. (1998) are critiqued here. Two hundred and fifty-four papers were identified that related to some aspect of PA promotion in primary care. Of these, twelve satisfied the inclusion criteria outlined above. Forty five sets of evaluation data from schemes were made available to Riddoch et al. (1998).

Riddoch et al. (1998) determined from a review of the data that the research methodology was consistently flawed and that the outcomes reported were often in direct contrast to those reported in the literature. Furthermore, few schemes reported the use of any psychological model of behaviour change and in the rare instance that schemes reported the integration of such a model, the details were not made available. Case study data revealed that common factors that appeared to consistently contribute to the success of schemes included: staff enthusiasm, working within alliances and maintaining good communications between organisations, designing individual exercise programmes that are tailored to each patient's needs, individual supervision, and having a low-cost policy, especially in areas of low income.

Of particular interest is that Riddoch et al. (1998) reported that being part of a scheme for psycho-social reasons was predominantly important for participants and that even small achievements were perceived as meaningful. It was indicated that

those referred found support, improvements in social activities and self-confidence. Moreover, Riddoch et al. (1998) reported that patients experienced an improved QoL and that those who suffered from anxiety and depression were particularly seen to benefit. However, it is unclear just how many of those referred for exercise experienced an improved QoL and the magnitude of these perceived changes. This is perhaps because the authors stated that only three of the schemes included in the review reported the use of recognised research tools (e.g., Short Form 36; Brazier et al., 1992) and that many schemes reported that the data generated from such measures was unavailable. Riddoch et al. (1988) claimed that there were difficulties with the transfer of data between those professionals involved in the delivery of exercise interventions and members of the research team. This is perhaps the result of the lack of consistent measurement tools available to practitioners with which to share and report such data. What may be beneficial to improving the evaluation of exercise referral schemes more generally would be to adopt a universally agreed evaluation criterion set such as the approach employed by weight management interventions (e.g., National Obesity Observatory, 2009).

Riddoch and colleagues (1998) suggest that the discrepancy between the case study data and published studies, in terms of the scale of impact of a PA intervention upon PA level and other reported parameters, would be improved by gathering quality data across schemes using valid measures. Indeed the authors proposed that this would provide substantial evidence related to effectiveness. Finally, Riddoch et al. (1998) highlighted that there is no standard recommendation for theoretically driven behaviour change strategies for exercise schemes in the UK. However, the authors recommend relevant training in motivational interviewing (MI) for primary care and referral staff in order to maximise participant motivation. Since the Riddoch et al. (1998) publication, guidance regarding behaviour change at population, community and individual levels has been published (NICE, 2007). Although this guidance goes some way to raising the profile of the importance of health behaviour change at these levels, the document makes no reference as to the most appropriate *method* of delivery. Riddoch and colleagues (1998) suggest that MI may be the most appropriate behaviour change method. Furthermore, recent developments within the arena of PA have demonstrated favourable results with respect to the feasibility of MI within primary care (e.g., Bull & Milton, 2010; Hilton, Milton, &, Bull, 2009). This is particularly poignant as the history of MI lies within the domain of addictive

behaviours (Hilton & Poulter, 2009) and the transferability of the method to PA behaviour is a relatively new concept.

More recently, and specifically related to exercise referral schemes in the UK, Williams, Hendry, France, Lewis, and Wilkinson (2007) conducted a systematic review to determine the effectiveness of exercise referral schemes to promote PA in adults. Eighteen studies were reviewed, including six RCTs, one non-randomised controlled study, four observational studies, six process evaluations and one qualitative study. In addition, two of the RCTs and two of the process evaluations incorporated a qualitative component. Williams et al. (2007) drew upon the recommendations outlined in the NICE (2006c) review and Morgan (2005; reviewed below), and concluded that based on the current data available, exercise referral schemes increased PA in some people, but that it was more costly than usual care. Morgan (2005) accounts for this claim by proposing that PA initiated by a referral into exercise may not be maintained in the long-term and that attendance was poor in the studies reviewed. Attrition has clear implications for the cost-effectiveness of such schemes and the low levels of long-term adherence supports the findings of the NICE guidance (2006c). Furthermore, low attrition highlights the necessity to address the current psychological evaluation shortfall experienced by schemes because it is reasonable to propose that psychological processes play a key role in encouraging long-term behaviour change.

Indeed, Williams et al. (2007) reported that the main challenges for future schemes are to increase adherence and improve uptake and they suggest that this may be achieved through considering readiness to engage in behavioural change, or by considering individual differences in self-determination and behavioural regulation. It is also suggested that exercise programmes should be more closely tailored to individual preference and that barriers identified by qualitative studies could be addressed.

The qualitative studies reported in Williams et al.'s (2007) review identified that participants had derived psychological, physical and social benefits as a result of attending the schemes. Dissatisfaction related to inconvenient operating hours for working people, congested facilities, insufficient staff, intimidating gym environment or equipment, narrow range of activities, and limited social interaction. Reasons for non-adherence included lack of self-efficacy and poor body image; poor organisation

of the scheme, such as inconvenient opening hours or inadequate supervision; poor personal organisation, such as finding time, transport, or interruptions of routine by illness or holidays; adverse social or psychological factors, such as poor social support, feeling uncomfortable in the gym environment; and an exercise leader lacking motivational skills.

Alongside more comprehensive and inclusive reviews of exercise interventions, including exercise referral, Morgan (2005) adopted a more limited outcome focus approach and investigated the capacity of exercise referral schemes to increase physical activity levels only. Nine studies met the inclusion criteria for review, four from the UK, four from the USA and one from New Zealand. Studies were reviewed if they met the following criteria: a) interventions providing access to exercise activities and/or facilities, b) studies that were experimental or quasi-experimental, c) studies that included a control group, d) interventions based in primary care settings and e) interventions that included an exercise component with measures of PA level or adherence. The review indicated that slightly active older adults and those who are overweight but not obese were most likely to benefit from an increase in PA as a result of participation in an exercise referral scheme. However, as the majority of studies and reviews conclude, these increases may not be sustained over time and Morgan urges for the identification of strategies to increase long-term adherence. The notion of long-term adherence, lack of longitudinal data for the evaluation of exercise referral schemes and the need for motivational strategies are common issues that run throughout the literature.

Utilising a randomised controlled trial (RCT) study design, Taylor, Doust, & Webborn (1998) conducted an evaluation study that examined the effects of a typical UK exercise referral scheme on modifiable coronary heart disease risk factors. Some positive health changes were noted, but the most significant improvements were observed for those who had attended 15 or more exercise sessions of a 10-week programme (high adherers). Three-hundred and eighty-nine smokers, hypertensive or overweight individuals were referred and subject to anthropometric and blood pressure measures at baseline, 16, 26 and 37 weeks. One of the strengths of the study was that, in contrast to the self-report method, data regarding adherence to the programme was collected from computerised leisure centre records. Reduction in skin-folds were noted up to 26 weeks in high adherers compared with controls, and a

reduction in systolic blood pressure was maintained up to 37 weeks among high adherers but only in comparison with low adherers.

Taylor and colleagues (1998) also recognised the challenges of implementing a RCT to evaluate health services research; namely, the self-selection bias of sampling procedures – smokers were less likely to enter the study while overweight individuals were more likely to participate. Similarly, those who were initially more active were more likely to complete the study. Opportunistic recruitment methods rather than mailed invitations coupled with motivational strategies to remain in the study were suggested for potential improvements for the study design. Indeed, integrating motivational strategies into routine exercise referral practice can only help to support those referred through the behaviour change process in addition to facilitating research/practitioner collaborations and maintaining the integrity of research designs. More traditional approaches to increasing PA such as the exercise referral scheme have been limited in this respect. However, as previously mentioned, recent recommendations regarding the use of brief interventions with respect to PA (NICE, 2006a), coupled with early indications of the feasibility of MI as an approach to supporting the PA behaviour change process (e.g., Bull & Milton, 2010; Hilton, Milton, &, Bull, 2009), may well assist in remedying this limitation because MI has the capacity to support long-term behaviour change.

While studies such as Taylor et al. (1998) focussed on the impact of the exercise referral scheme in isolation, Isaacs et al. (2007) conducted one of the most comprehensive investigations of the impact of exercise referral upon a number of parameters including a cost analysis, but as a comparison to community-based walking and advice only. Participants (N = 943) aged between 40 and 74 years who had at least one coronary risk factor and who were inactive were included in the study. Participants were randomised to one of three groups: a 10-week programme of supervised exercise classes, two to three times a week in a local leisure centre; a 10-week instructor-led walking programme, two to three times a week; an advice-only control group who received tailored advice and information on physical activity including information on local exercise facilities. At the end of the final assessment, all control group participants were offered a referral to the leisure centre. This approach to study protocols is typical of RCTs where it is considered that control group participants may not have the same opportunity for health improvement as those in the experimental group.

Primary outcome measures included changes in self-reported exercise behaviour, blood pressure, total cholesterol and lipid subfractions. Secondary outcomes included changes in anthropometry, cardiorespiratory fitness, flexibility, strength and power, self-reported lifestyle behaviour, general and psychological health status, quality of life and health service usage. Isaacs et al. (2007) documented that participants' indicated that the cost of leisure services, the type and duration of support received from leisure staff the time available to exercise at the leisure centre and pain management were commonly reported important considerations for participants.

An economic analysis of the cost of providing and using the services was also provided which, in addition to the psycho-physiological outcome measure design should be considered a strength of the study. If PA is to be considered a feasible prescription for health, as the evidence base grows, the necessity to provide data that allows for cost comparisons with usual healthcare practice (e.g., prescription costs) will be required. UK healthcare is under increasing pressure to demonstrate cost-effectiveness.

To enhance the direct relevance of the data generated by Isaacs et al. (2007) for the current thesis, only the psychological data reported by Isaacs et al. (2007) will be presented here. Six months after beginning the trial, all three groups showed improvement in anxiety and mental well-being scores. Although there were no differences between groups, the leisure centre exercise referral group and walking group maintained this improvement at one year. Measures of anxiety and depression were assessed using the Hospital Anxiety and Depression Scale (HADS; Snaith, 2003). Although the HADS has received fairly consistent support in terms of psychometric properties (e.g., Mykletun, Stordal, & Dahl, 2001), the measure is only capable of detecting the more negative aspects of mental health (i.e., anxiety and depression). This is rather limited given the documented potential impact PA may have on positive mental health and well-being (e.g., DH, 2004b; Fox, 1999). A generic QoL measure – the Short Form 36 (SF-36; Brazier et al., 1992) was also administered to participants, perhaps to broaden the assessment limitations of the HADS and in particular Isaacs et al. (2007) report more thoroughly on the mental health cluster of scales. A between-groups, one-way analysis of variance (ANOVA) revealed a significant improvement in mental health scores between baseline and six months for all three treatment groups $F, (1,315) = 23.02$, $\underline{p} < .05$. There were no

significant differences in depression scores between treatment groups at either baseline or six-month follow-up. However, a significant difference in depression scores for the leisure centre group between baseline and six-months was observed F, (1,176) = 16.98, $\underline{p}$ <.05. These findings in particular further strengthen the rationale for addressing the psychological outcomes of exercise referral within structured exercise settings.

Measures of attitudes and intention to exercise that were assessed in this study were based on the Transtheoretical Model (TTM; Prochaska, DiClemente & Norcross, 1992). The Stages of Change component of the model was measured using a single item statement designed to represent the five stages of change. Participants were asked to respond to the single item measure at each assessment stage. It was noted that mean stage of change progressively improved across the periods of assessment and further analysis (Bonferroni- adjusted *t*-tests) showed no difference between groups at baseline.

At 10 weeks the leisure centre and walking groups demonstrated a significantly higher mean stage of change than the advice-only control group, but did not differ from each other. Mean stage of change was also significantly higher for these groups compared with the baseline measures. At six months, mean stage of change remained significantly higher for both leisure and walking groups compared with controls. However, at this assessment point walkers also had a significantly higher stage of change than the leisure centre group. An extension of this analysis to one year for the leisure centre and walking groups, both retained a significant improvement in mean stage of change compared with baseline and the mean stage of change remained higher for the walking group than for the leisure centre group at the one-year follow-up stage.

Furthermore, Isaacs et al. (2007) concluded that supervised leisure centre-based exercise classes and instructor-led walks were no more effective than advice-only in initiating and sustaining increased physical activity, and in modifying cardiovascular risk factors. However, psychological parameters notably improved to a greater extent in the walking and leisure centre groups. Isaacs et al. (2007) draw upon the NICE (2006a) recommendations regarding brief interventions and suggest that a combination of a brief intervention and an exercise referral design may be beneficial for improved participant outcomes. As mentioned previously, such work has already

begun in the form of the recently developed Let's Get Moving programme process (e.g., Bull & Milton, 2010; Hilton, et al., 2009). However, this work is in its infancy and greater insights are needed into the outcomes of this approach and MI treatment fidelity in particular (Bull & Milton, 2010).

Effective methods of maintaining motivation to exercise within clinical settings also have a role in helping individuals to initiate health behaviour change. James et al. (2008) sought to investigate those factors associated with referral uptake and participation and although the study was largely demographically focussed, the outcomes and recommendations made by the authors support the necessity for the inclusion of the evaluation of psychological responses to exercise referral. Over a three-year period (N = 3,762) participants were recruited and the data used for the study capitalised on that which was routinely collected from a county-wide physical activity referral scheme which should be considered a strength of the study design in terms of transferability of results and relevance for routine practice.

The results of the study indicated that the scheme coordinator was less likely to make contact with participants with a referral for overweight/obesity (odds ratio 0.59, 95% confidence interval 0.36 – 0.95, p = 0.03) or a mental health condition (0.35, 0.19 – 0.66, p = 0.001) than those referred for a cardiovascular condition. The reason a participant had been referred into the scheme and whether they were male or female were both associated with being assigned to a leisure provider (p = 0.047). Compared with patients with a referral for a cardiovascular condition, patients with a referral for overweight/obesity (p = 0.035) or a mental health condition (p = 0.016) were less likely to be assigned to a leisure provider. Uptake of referral was associated with the reason for referral and the referring health professional. Furthermore, those referred for overweight/obesity ($\underline{p} < 0.001$), musculoskeletal health (p = 0.042), mental health conditions ($\underline{p} < 0.001$), and 'other' reasons (p = 0.003) were less likely to take up a referral opportunity than patients with a referral for a cardiovascular condition. Compared with patients referred by their GP, those referred by another health professional (e.g., dieticians, psychiatrists) were less likely to take up a referral opportunity (p = 0.002). This is contrary to the findings of Dugdill et al. (2005) who found that patients were more likely to adhere to a programme if referred by an allied health professional, for example a Practice Nurse. Perhaps further work needs to be undertaken to investigate this discrepancy further. James et al. (2008) also found that

females were less likely to complete a referral ($p = 0.043$), whereas increasing age was positively associated with completion ($p < 0.001$).

It is likely that those referred for a cardiovascular condition were more frequently assigned to a leisure provider because of issues related to clinical governance and overall responsibility of those diagnosed with a cardiac condition who are often referred to as high-risk populations (DH, 2002). It is also likely that increased age was positively associated with scheme completion because those of retirement age would not have to manage work-related time barriers to PA participation which has been previously reported in the literature (Biddle & Mutrie, 2001).

Based upon the findings, James et al. (2008) make three recommendations for practice and one for research. For practice, the authors indicated that: better targeting and an exploration of exercise barriers pre-referral by health professionals could improve uptake; a greater choice of activities other than those that are facility-based may assist with continued participation; and schemes should allow for a greater amount of flexibility for those individuals who experience genuine time constraints or for those who would benefit from a drop-in approach due to fluctuating ill health symptoms. In particular, James et al. (2008) recommend this approach for those who have been referred for mental ill health. Future research considerations urge for a continuation of investigations that seek to determine the efficacy of referral schemes and that qualitative methods would assist with developing an understanding of why certain groups (e.g., young adults, women, those who are overweight/obese or mental health conditions) may be less suited to such schemes.

Attempts to evaluate exercise referral schemes using a qualitative approach have been previously published. Singh (1997) interviewed 13 participants between 30 and 61 years of age. All but two were female and attending a referral scheme at a local leisure centre in London, UK. The results reported are limited in that the method(s) of analysis of the data are not reported and only two emergent themes are described (i.e., patients and their bodies and patients and their beliefs). It is uncommon for so few themes to be reported without some description of how the data may have been grouped together into ‘higher order themes’ as a result of thematic analysis.

Consequently, some of the verbatim quote examples provided by Singh (1997) reflect missed opportunities to code the data more thoroughly in order to represent

what was said by those interviewed more accurately. For example, a report given by a participant describes how exercise has helped to build up general confidence. However, because this data is grouped under the generic theme of patients and their bodies the opportunity to code this data as a psychological outcome has been missed. Similarly, in an additional verbatim quote example, the same participant elaborates upon the role of confidence and the importance of the resultant companionship with others participating in the scheme. However, because Singh (1997) reports this data under the generic theme of patients and their beliefs, again some of the specificity of reports made is lost.

With respect to qualitative analysis, and especially given the conflicting evidence base for exercise referral, it is critical that the level of analysis supports the deconstruction of these inconsistencies and fluctuations. Subsequent qualitative studies have assisted in this matter (e.g., Crone et al., 2005; Crone-Grant, 1999, 2001) and further investigations into the role of group versus individual exercise have helped to tease out the impact of these differing modes of exercise upon participant adherence and motivation. For example, Carron, Widmeyer, & Brawley (1988) demonstrated that group cohesiveness was related to individual adherence behaviour across different exercise types including fitness classes, recreational team sports, and elite team sports. Booth, Bauman, Own and Gore (1997) utilised a survey approach to ascertain that, for inactive Australian residents, a group mode of exercise would be most preferable for individuals between the ages of 18 to 39. However, Dugdill et al. (2005) have subsequently reported that individuals within the age range of 18 – 30 were the most likely not to attend or contact an exercise scheme following referral. There are inherent difficulties with establishing the level of accuracy of self-report data, although this method is often considered the most feasible approach to data collection within settings such as exercise referral schemes (e.g., Isaacs et al., 2007).

Exercise referral schemes often refer individuals into either group or individual opportunities for PA. The benefits of social support and interaction have been identified as important factors for those referred (e.g., Crone et al., 2005; Hardcastle & Taylor, 2001) which may indicate that for the purposes of longer-term adherence to exercise and the associated benefits of improved life-quality, group-based exercise may be more beneficial than individual-level PA. However, some of the more well-

established exercise referral schemes in the UK honour that those referred are more likely to adhere to an activity that they enjoy and therefore offer a range of exercise options. These options often include both group and individual exercise and to help establish with greater certainty whether group-based exercise is consistently more effective than individual-level PA at enhancing adherence and the mental and physical benefits of PA; further research in this respect is warranted.

In summary, the efficacy of exercise referral schemes has been given a fair level of attention in the literature (e.g., Dugdill et al., 2005; Harrison et al., 2004; Morgan, 2004; Riddoch et al., 1998) However, significantly more attention is given to physiological outcome measures, despite these measures being the least likely to change during an initial 12-week exposure to structured exercise (e.g., Dugdill et al., 2005). Relevant psychological and environmental parameters tend to be ignored (Crone, Smith, & Gough, 2005; Dugdill et al., 2005) and the evidence base to support the efficacy of exercise referral schemes is often regarded as conflicting or inconclusive (Crone et al., 2004; Graham et al., 2005).

Perhaps particularly in terms of increasing PA levels in the longer-term, the evidence base for the impact of exercise referral is not convincing (NICE, 2006a; 2006c) and potential stakeholders are in a position of uncertainty as to the value of an investment in such an intervention. However, this uncertainty is primarily based upon the potential of schemes to impact upon PA level only and excludes other health parameters (NICE, 2006a; 2006c). Despite the recognition of the potential for a referral into exercise to impact positively upon psychological parameters (DH, 2002), the evidence base is unclear. This is largely due to the medical-model driven physiological evaluation bias of health services and the lack of psychologically focused scheme delivery and evaluation tools. Such shortfalls make it difficult to examine the processes of attitudinal shift, behaviour change and perceptions of exercise-related life-quality.

1.5 Moving Towards the Holistic Assessment of Exercise Referral

Research has examined the potential of PA to impact positively upon a range of psychological outcomes including mood (e.g., Biddle, 2000), self-esteem (e.g., Fox, 2000), anxiety (e.g., Taylor, 2000) and depression (e.g., Mutrie, 2000). It is beyond the scope of this thesis to review each of these outcomes in depth. Therefore, this review will provide relevant background as to why psychological outcomes are an important consideration for exercise referral schemes, and why a measure of QoL may be the most feasible measurement option for exercise and health practitioners.

In light of the NQAF document (DH, 2002), there is an acknowledgement that psychological measurement change may be more likely than physical for exercise referral participants – even during the short course of a 10-week programme. Additionally, explorations of the link between exercise/PA, psychological well-being and mental health lend support to the notion of exercise and PA as a valuable treatment option for ill health, and as a method of enhancing positive mental health and well-being (Fox, 1999). However, in practice minimum attention is given to our understanding of the mental and social well-being of individuals who often have varying degrees of chronic ill-health, and are referred for exercise. This is despite recommendations to refer individuals for PA to improve their mental health (Edwards, Ngcobo, Edwards, & Palavar, 2005).

Edwards et al. (2005) provide evidence for the potential impact of PA upon psychological parameters when they reported that participants who engaged in regular PA perceived themselves as having more autonomy, personal growth, environmental mastery, purpose in life, positive relations with others, self acceptance, sport competence and conditioning than non-exercisers. The authors present their findings as a basis for encouraging health professionals to routinely consider the referral of individuals with mental health and/or stress-related problems to health clubs. Furthermore, these findings also demonstrate why it is important for exercise professionals to consider a number of physio-psychological outcomes. Just as there has historically been a limited physiological evaluation bias of referral schemes, there may be a temptation to replicate this limitation to few psychological parameters. The introduction of an exercise referral sensitive QoL measure

developed directly from the reports of participants would reduce the likelihood of this limited approach to evaluation.

There are few studies that have sought to specifically identify the mental health outcomes for exercise referral participants. Crone et al. (2005) have demonstrated a psycho-social explanation of the relationship between PA and mental health for exercise referral participants. The explanation of self-acceptance and the importance and interrelationship of social support and the physical environment were found to be important for increasing the likelihood of positive experiences of a referral into exercise. This work coupled with previous qualitative explorations of the experiences of exercise referral participants (e.g., Crone-Grant & Smith, 1999; 2001) has prompted interest in establishing the wider outcomes and experiences of referral into exercise beyond physical adaptation, exercise adherence and disease management. Prior qualitative investigations have indicated the need for social support and the cultural scene (Crone-Grant & Smith, 1999) and more recently, social support and networks, and a feeling of competence with the equipment and the environment as important issues for exercise referral participants (Crone-Grant & Smith, 2001).

Considerations of the impact of PA and exercise upon parameters other than physiological indicators often fall into investigations of perceived well-being as a primary or secondary outcome. Bloodworth and McNamee (2007) provide a philosophical critique of conceptions of well-being in psychology and exercise psychology research. Following a comprehensive exploration of subjective and objective well-being and the notion of psychological affect, Bloodworth and McNamee (2007) assert that the potential that physical activity has to increase our opportunities, not only to conduct daily tasks with greater ease but to learn new skills, get out of the house, meet new people, try new things, challenge ourselves and accomplish feats, cannot be reduced solely to affect or satisfaction.

In this way, while studies of the relationship between exercise, PA and affect or satisfaction for example, may inform psychological relationships at the micro level (i.e., quite focussed and specific), providing the tools for practitioners to explore the relationship between exercise, PA and QoL will help to add to our understanding of these relationships at more of a macro level (i.e., much more broad and large in scale). Therefore, approaching the evaluation of exercise initiatives from a QoL perspective will not only shift outcome parameters beyond a physiological domain

but facilitate the quality and depth of psychological investigations to include a variety of psycho-social factors inclusive of and in addition to those of affect and satisfaction.

Polzien, Jakicic, Otto, Winters-Hart, Thomas and Garda (2005) support the value of psychological data collection by reporting that for weight-loss participants enrolled in an exercise program, greater *perceived* (author's emphasis) benefits may enhance exercise adoption. Therefore, it may be argued that, particularly in terms of facilitating behaviour change, the identification of the benefits of PA perceived by participants should be encouraged, in addition to or even in the absence of physiological change. The identification and amplification of these individual perceptions could have a valuable role to play in encouraging initial attitudinal shifts as to the benefits of PA and subsequent behaviour change. The communication of wider exercise outcomes is missed by the employment of prescriptive physiological outcome measures (Dugdill et al., 2005), therefore a focus upon more individual, person-centred methods of engaging the less active should be encouraged. Anecdotal reports from exercise referral schemes that exemplify diverse outcomes via case study reports have given rise to an interest in quantifying these reports into a common, collective language that is easily communicated within and between exercise schemes. Enabling health and exercise professionals to communicate data in such a way will also improve the ability to share patient outcomes with the referring practitioner more easily which has been a limitation of exercise referral schemes previously identified (Graham, Dugdill, & Cable, 2005).

The importance of PA for both quality and quantity of life has been emphasised and the continued development of appropriate measures to capture this has been encouraged (Dubbert, 2002). In practice and research, one of the most common attempts to capture more psycho-social type data in addition to that generated from physiological measures has been to introduce generic QoL measures into evaluation strategies. Until an exercise referral specific measure becomes available, generic measures have been the only tools available to practitioners and researchers alike. Given the convincing relationship between life quality and exercise behaviour for clinical populations (e.g., Motl & Snook, 2008) coupled with an increasing need to evaluate exercise referral schemes more thoroughly (NICE, 2006a), holistically

(Dugdill, et al., 2005) and with a particular emphasis upon psychological parameters (DH, 2002) the necessity for appropriate tools to be developed is evident.

1.6 The Conceptual and Assessment Challenges of Life-Quality

The notion of life-quality has the potential to broaden more traditional healthcare perspectives that have tended to focus upon disease and illness. Understanding life-quality usually takes the form of the development of conceptual models. However, there is no definitive conceptual model of QoL nor is a single model or assessment tool utilised consistently within research or practice settings. Although not a model of QoL as such, the proposal of a bio-psycho-social model of health assisted in addressing the psychological, social, and behavioural dimensions of illness (Engel, 1977) and in this respect is worth acknowledging as a contribution to the development of healthcare systems that include more holistic patient outcome measurement.

The historical development of models of life-quality reflects the health psychology disciplinary perspective of the current thesis. For example, Brown, Bowling & Flynn (2004) identified that QoL models have ranged from needs based approaches such as Maslow's (1954; 1968) hierarchy of human needs, to classic models based on psychological well-being, happiness, morale, life satisfaction (e.g., Andrews & Withey 1976; Larson 1978), social expectations (e.g., Calman 1983) or the individual's unique perceptions (e.g., O'Boyle 1997). More recently, Wilson and Cleary (1995) have argued that biological and physiological factors, symptoms, functioning, general health perceptions, and overall quality of life are the most relevant indicators of a conceptual model of QoL. While Ferrans (1996) proposed a QoL model comprising four differing conceptual domains: health and functioning, psychological/spiritual, social and economic, and family. The diversity of these models indicates that QoL is an inherently dynamic and complex concept which proves problematic for measurement (Brown et al., 2004).

Nevertheless, QoL and health related quality of life (HRQoL) measures are accepted as part of routine practice and are often integral to clinical research trials. However, the literature demonstrates two key challenges regarding QoL assessment - namely, the definition (e.g., Cohen, Mount, & MacDonald, 1996) and the measurement of the

concept (e.g., Allison, Locker, & Feine, 1997; Holmes, 2005). Contributions to attempts to define QoL have, in part been provided by the development of conceptual models as explained above. From an applied perspective, exercise referral personnel are left wanting when it comes to a measure of life-quality specifically designed and validated for the collection of routine psychological response evaluation data. Despite this shortfall, it is important to acknowledge some of the more conceptual and philosophical issues related to QoL measurement although these are critiqued more fully in the final Discussion and Conclusions of the current thesis (Chapter VII).

Quality of life studies and measurements have been described as a method to prevent the separation of a patient's body from a patient's biography during delivery of care (Roy, 1992). Despite there being no 'gold standard' measure of life-quality (Holmes, 2005; Pais-Ribeiro, 2004), the inclusion of measures of QoL into healthcare systems are representative of the necessity to address those under care as an integrated whole rather than a series of disconnected parts or disease states (Holmes, 2005). Although it is generally accepted that QoL is a valuable and necessary outcome measure within healthcare settings, the subjective, transient nature of QoL has led to considerations regarding the efficacy of accurate measurement.

HRQoL has been described as participant's perceptions of function (Rejeski, Brawley, & Shumaker, 1996). Consequently, health-related quality of life measures have been developed to describe aspects of an individual's subjective experience that relate both directly and indirectly to health, disease, disability, and impairment (Carr, Gibson, & Robinson, 2001). This may account for the wealth of HRQoL measures that are applied to the assessment of exercise upon a given medical condition – the focus of which are usually upon patient perception's of function. For example, QoL tools have been developed for the assessment of the following: ankylosing spondylitis (Doward et al., 2003), eating disorders (Las Hayas et al., 2006), menopausal women (Hilditch et al., 1996) and obesity (Kolotkin, Crosby, Kosloski, & Williams, 2001) to name a few. The very nature of HRQoL tools means that QoL is predominantly assessed in terms of illness and morbidity. More positive dimensions of life quality such as social relationships, for example are often omitted from such measures. Furthermore, these tools are usually administered as part of a clinical trial or controlled study and are limited in the range of questions that are

meaningful to exercise referral participants. As yet, a measurement tool that is designed specifically for exercise referral is lacking.

For both the participant and the exercise referral practitioner, it is often the exercise outcomes that reflect changes in mental health, active daily living and optimum functioning and well-being that are most responsive to a dose of exercise, easily recognised by the participant, often considered the greatest achievement, and consequently, most meaningful. The latter of these outcomes (optimum functioning and well-being) has been alluded to as a definition of QoL (Cella, 1994). Maintaining or improving the quality of life has even been referred to as "the essence of healthcare" (Holmes, 2005, p. 493). Furthermore, it has been proposed that the essential goal of good health is to improve the *quality* of years an individual lives, and not necessarily to add years to life (Chandra, 2001). It is for these reasons and the recognition that participant responses to exercise referral often go beyond physiological outcomes that exercise professionals have attempted to capture some of this valuable data.

A popular method has been via the administration of generic QoL health assessment tools. In practice, one of the most commonly used measures is the EuroQol (EQ-5D; Brooks, 1996) while for the purposes of research the Short-Form SF-36 (Brazier et al. 1992) is often the instrument of choice. The EQ-5D is a short self-administered questionnaire consisting of three questions each representing five dimensions of health. These dimensions are mobility, self-care, usual activities, pain/discomfort and anxiety/depression. The measure also includes a generic scaling question regarding current perceived health status. The SF-36 consists of eight scaled scores that represent the sums of the questions in each section. As the name of the measure would suggest, the 36 questions in total are transformed into a 0-100 scoring scale on the assumption that each question carries equal weight. The eight dimensions of the SF-36 include: vitality, physical functioning, bodily pain, general health perceptions, physical role functioning, emotional role functioning, social role functioning and mental health. It is the comprehensive nature of these measurement dimensions perhaps that reflects the appeal of the SF-36 for research purposes. However, the measure does not translate well into practice because of the length of time respondents take to complete the 36 questions and what for time-restricted health and exercise practitioners is often perceived as quite a complex scoring protocol.

In an attempt to reduce respondent burden and increase the ease of which the Short Form is administered in practice, the Short Form (SF-12; Ware, Kosinski, & Keller, 1996) was developed. The SF-12 measures the same eight dimensions as the SF-36 but utilising fewer questions. Although a shorter version of the popular SF-36 would seem more appropriate for use within exercise settings, a lack of consensus regarding the QoL measure most suitable for exercise referral may account for the lack of its use. This is despite recent support for the use of the SF-12 for use with obese and non-obese primary care patients. Particularly in terms of detecting QoL differences associated with body mass index (Wee, Davis, & Hamel, 2008). Similarly, the WHOQOL-BREF (Skevington, Lofty, & O'Connell, 2004) offers a relatively short (26-item) generic measure of life-quality, in the main the face-value of questions indicate suitability for use within exercise referral settings and the WHOQOL-BREF is reasonably simple to score yet the measure has not been integrated into national evaluation procedures.

Although these measures go some way to capturing valuable QoL data, the difficulty still remains that the generic items do not reflect those specific and common responses to and outcomes of exercise referral as identified by those participating. Consequently, those aspects of QoL that are important to exercise referral participants are under-represented in the literature and in practice.

Within exercise referral settings, generally, it is felt that the current and available QoL tools (e.g., EQ-5D; Brooks, 1996; SF-36; Brazier et al., 1992) are too lengthy, difficult to score, and the questions are perceived as irrelevant to exercise referral populations, too intrusive or inappropriate to be administered by PA practitioners. They are often used as a compromise to attempt to address the wider responses to and outcomes of exercise referral because they are all that is available. However, the use of measures such as the SF-36 within clinical exercise settings has indicated that improvements in perceived QoL can be observed at baseline to three months and one year (Alexander & Wagner, 2006). Therefore, although the current and available tools with which to assess QoL may not be suited to the exercise referral environment, the justification for assessing QoL as a valuable outcome of a referral into exercise is evident.

Driven by the observation that there is no consensus on the definition or measurement of QoL (Skevington, Lofty, & O'Connell, 2004), Moons, Budts, and

Geest (2006) provided a critical review of the conceptualisation of QoL. Understandings of QoL in the biomedical and nursing literature are described as: a) normal life, b) social utility, c) utility, d) happiness/affect, e) satisfaction with life, f) satisfaction with specific domains, g) achievement of personal goals and h) natural capacity. An evaluation of the different concepts in the context of the inherent problems created by each principle gave rise to Geest (2006) concluding that satisfaction with life is the most suitable approach to defining QoL. However, satisfaction with life still represents a subjective term that would be difficult to quantify for practitioners working within the settings targeted by Moons et al.'s (2006) review. Perhaps it is difficult to define quality of life in a generic sense and more appropriate to consider the significance of the term for specific populations who share common lifestyle and health circumstances. While Moons and colleagues (2006) went some way to considering conceptions of life-quality within the biomedical and nursing literature, practitioners working within such settings would require further guidance regarding those constructs that are most suitable to represent satisfaction with life and that would indicate degrees of satisfaction with these relevant constructs specifically.

One of the most concise definitions of QoL has been offered by Cohen et al. (1996) who proposed that QoL simply describes subjective well-being. The term subjective well-being is a term in itself that is problematic because of the diversity of its interpretation. However, Diener, (2000) has attempted to capture the terms complexity by suggesting that it represents people's evaluations of their lives that are both affective and cognitive. Cohen et al. (1996) echo previous proposals of the subjective nature of QoL (Aaronson, 1990; Campbell, 1976; Cohen, & Mount, 1992; Guyatt, Feeny, & Patrick, 1993) and in that respect ascertain that a reliable level of measurement can only be established by considering the person whose QoL is being evaluated. The authors also recognised that the importance of life-quality is likely to vary over time for individuals, despite experiencing the same objective circumstances, and that we do not have a clear and complete working understanding of what constitutes QoL. Suggestions to improve this understanding, and in particular to support patients better, are made by proposing that those cared for are well placed to identify the most suitable and relevant domains to be measured. Once appropriate levels of reliability and validity are established, Cohen et al. (1996) proposed that attention should be given to the integration of these domains into healthcare in such a way that QoL is considered in its entirety – in view of the 'whole person.'

Furthermore, the authors stipulate the importance of developing measures that are transferrable from research to practice so as to facilitate use in clinical settings.

Hunt (1997) clearly presents the QoL measurement difficulties experienced by clinicians working within the healthcare arena. Hunt (1997) asserted that the varying content of questionnaires that imply divisions in terms of definition and inconsistencies that would not be tolerated for medical terminology, causes frustration and confusion for practitioners. The notion of consistency regarding defining QoL, particularly where health evaluation is concerned is deemed as paramount when such assessments contribute to medical care and influence the lives of patients. Indeed, given the lack of definition and measurement consensus, the inclusion of measures of life-quality in clinical settings is deemed as unethical by Hunt (1997).

There are a wealth of generic and health-related QoL measures available to practitioners and researchers alike. However, a QoL measure specifically sensitive to those items of importance for exercise referral populations has yet to be developed. Some of the variance in the item-content of measures designed to assess the life-quality of individuals diagnosed with differing disease may be accounted for by the sheer volume of measures available for a given health condition. While in theory exercise referral professionals could select those disease-specific measures that have been designed and validated appropriately and use them to assess the life-quality related to each health condition with which a person has been referred, the number of questionnaires that would be required is just not feasible. Rather, an exercise referral sensitive QoL measure that has been designed by a sample of participants with health conditions and demographics representative of those referred to schemes is likely to capture those shared QoL related items of importance. Furthermore, as there have been no previous measures designed for this use, the item conflict and inconsistency issues between measures designed for the same purpose and highlighted by Hunt (1997) have been avoided.

The majority of philosophical and practical challenges of QoL measurement that have been presented in the literature have focussed on the general concept and between-person differences in determining instrument content. However, Allison, Locker, and Feine (1997) have addressed the possibility of within-person differences

(i.e., an individual changing the standards by which they assess their QoL and the subsequent effects upon valid measurement) creating challenges to QoL assessment. Despite the difficulties of changes in the point of reference from which life-quality is assessed, in other words the possibility of attitude change between assessment points, Allison and colleagues support the use of QoL measures within health care settings. Indeed the authors call for the development of valid and appropriate measures and assert that such measures make an important contribution to health care evaluation.

1.7 The Exercise and Quality of Life Relationship

Perhaps as Allison et al. (1997) recognised, although debate regarding the definition and measurement of life-quality continues, the necessity to capture indicators of mental and physical well-being within healthcare settings is such that it is sufficient to take precedence over these philosophically focused debates. Crone et al. (2005) recognised that 'mental health' is more commonly researched and defined in terms of disease states (e.g., depression), rather than having a positive dimension (i.e., well-being). Those studies that have explored the link between PA/exercise and disease states should, therefore, perhaps be more accurately referred to as studies that explore the link between mental *ill* health and PA/exercise. However, one method of addressing the relationship between PA/exercise and mental health in terms of the more complete bipolar understanding of mental health (i.e., taking into account mental health and illness) that together inform a complete state of mental health (Keyes, 2005) is via the use of quality of life (QoL) measures. Indeed, the measure of life-quality has been deemed as the primary end-point in clinical settings (Pais-Ribeiro, 2004).

The exercise-QoL literature tends to focus on the relationship between exercise, perceived QoL and varied ill health conditions (e.g., Milne, Guilfoyle, Gordon, Wallman, & Courneya, 2007; Motl & Snook, 2008). Generally exercise is considered as a valuable opportunity to enhance perceived QoL for those who have been diagnosed with chronic or life-threatening illnesses that typically impact negatively upon perceived QoL. It is valuable to note that the ill health conditions that have been addressed in the literature in terms of their relationship with exercise and QoL are also typical of those conditions for which individuals are referred into exercise schemes (e.g., cancers, weight loss, multiple sclerosis and hypertension). Therefore,

it is reasonable to transfer the learning from the disease-specific relationships between PA and QoL into exercise referral directly.

The relationship between exercise and breast cancer is one of the most frequently reported (e.g., Burnham & Wilcox, 2002; Courneya, 2005; Courneya, Makey, Bell, Jones, Field, & Fairey, 2003; Pinto, Frierson, Rabin, Trunzo & Marcus, 2005). Support for the notion of including exercise as a valuable prescription for improving the QoL of breast cancer survivors is offered by Milne, Guilfoyle, Gordon, Wallman, and Courneya (2007). Breast cancer survivors (N = 289) responded to a questionnaire that assessed exercise attitudes, behaviour and perceived QoL. Additionally, participants responded to two open ended questions that aimed to explore perceptions of exercise and QoL throughout their experience of cancer and these responses were subject to inductive and deductive qualitative content analysis. Themes identified included: exercise behaviour, lifestyle, limitations and barriers, growth and priorities, and personal beliefs and values. Exercise was considered a valuable method of helping breast cancer patients to focus on long-term health goals, self-care and as an opportunity to engage in personal change. Individually tailored programs of exercise were recommended that take into consideration the likely barriers to exercise that may be experienced by this population. Such barriers include time, and issues related to health and age as well as access to suitable leisure facilities. Information regarding the benefits of exercise for recovery from breast cancer was encouraged, particularly to help alleviate any pre-exercise program anxiety.

The mode or type of exercise in terms of impact upon perceived QoL for breast cancer survivors has also been addressed. Ohira, Schmitz, Ahmed and Yee (2006) conducted a RCT study to determine if twice weekly weight training would impact upon perceptions of QoL and depressive symptoms as measured by the Cancer Rehabilitation Evaluation System Short Form (CARES-SF) (Schag, Ganz, & Heinrich, 1991) and the Centre for Epidemiologic Studies-Depression Scale (CES-D) (Radloff, 1977), respectively. A convenience sample of 86 participants were randomized into treatment (n = 43) and control (n = 43) groups. Thirty nine participants in the treatment group were followed up at six-months after completing nine common weight-training exercises for chest, back, shoulders, arms, buttocks, hips, and thighs twice-weekly.

Key findings included improvements in physical global QoL scores (CARES-SF) in the treatment group compared with the control group ($p = .006$). The psychosocial global score also improved significantly in the treatment group compared with the control group ($p = .02$). There were no changes in CES-D scores. Increases in upper body strength were correlated with improvements in physical global score ($r = 0.32$; $p < .01$) and psychosocial global score ($r = 0.30$; $p < .01$). Increases in lean mass were also correlated with improvements in physical global score ($r = 0.23$; $p < .05$) and psychosocial global score ($r = 0.24$; $p < .05$). The authors were led to conclude that this mode and frequency of training can improve QoL for breast cancer survivors, perhaps as a result of changes in body composition and strength. In this respect, current physiological measurement procedures within exercise referral settings may help to support the measurement of psychological parameters. There may be opportunities to consider correlations between changes in body composition and perceptions of life quality once the initial referral period has been completed and as part of a participant outcome review.

In addition to research conducted with breast cancer survivors, the specific relationship between exercise and the QoL of hypertensive patients has been investigated by Fernandez, Garcia, Alvarez, Giron, and Aguirre-Jamie (2007). Spanish hypertensive patients aged over 18 years ($N = 361$) were exposed to 'physical exercise' as directed by their GP. The authors refer to 'physical exercise' as physical training programmed by the doctor in accordance with the specific circumstances and capacities of each individual patient. Fernandez et al. (2007) clarify that for registered hypertensive patients, doctors should prescribe progressive aerobic physical exercise on at least three alternate days a week, with a duration of 30-60 minutes and at such intensity as to maintain the patient's pulse rate between 60% and 85% for their maximum heart rate (MHR, or 220 minus age in years).

Limitations of the study were recognised; in particular, the use of a cross-sectional analysis of a cohort which the authors identified may not be the best to determine causal relationships, there was a lack of control over factors such as the pharmacological treatments that the hypertensive participants may have taken, and there were challenges with controlling the degree of diabetes mellitus or dyslipidemia as possible co-morbidities within the hypertensive population that were included in the study. Additionally, the type, frequency, intensity and duration of the

exercise undertaken by participants was not controlled, monitored or reported in any way except for the hypertensive registry stipulation outlined above. However, the study reported that PA was associated with an improvement in all seven dimensions of life-quality for women and in five scales corresponding to psychological and physical dimensions for men as measured by the Spanish version of the Profile of the Quality of Life in the Chronically Ill Questionnaire (PECVEC) developed by the first author and colleagues (Fernandez-Lopez, Seigrist, Hernandez-Mejia, Broer, & Cueto-Espinar, 1994). Fernandez et al. (2007) concluded that GPs should routinely prescribe exercise for hypertensive patients with the stipulation that frequency, duration, intensity and evaluation should be included in the prescription.

Based on the premise that QoL is often compromised in those diagnosed with multiple sclerosis (MS), Motl and Snook (2008) explored the relationship between PA and QoL from a social-cognitive perspective driven by the hypothesis that the relationship between PA and QoL may be indirect and accounted for by self-efficacy. One hundred and thirty-three participants diagnosed with MS completed the Godin Leisure-Time Exercise Questionnaire (Godin & Shephard, 1985), Multiple Sclerosis Self-Efficacy Scale (Schwartz, Coulthard-Morris, Zeng & Retzlaff, 1996) and Multiple Sclerosis Impact Scale (Hobart, Lamping, Fitzpatrick, Riazi, & Thompson, 2001). Correlational analysis indicated that PA was associated with physical and psychological components of QoL. Furthermore, those who were more physically active had greater self-efficacy for function and control, and greater self-efficacy for function and control were associated with better physical and psychological components of QoL. It was suggested that PA be used as a method of decreasing the potential for reductions in perceived QoL, with a particular emphasis upon the role of PA in improving self-efficacy perceptions of those diagnosed with MS.

In contrast, an earlier study conducted by Romberg, Virtanen, and Ruutiainen (2005), concluded that exercise may improve functional impairment but not the life-quality of those diagnosed with MS. The RCT included 47 participants assigned to an exercise group in which they participated in a progressive resistance exercise for six months. The control group (n = 48) received no intervention. Participants were assessed at baseline and at six months using the Multiple Sclerosis Functional Composite (MSFC; Fischer, Rudick, Cutter, & Reingold, 1999), the Expanded Disability Status Scale (EDSS; Kurtzke, 1983) the Functional Independence Measure

(FIM; Stineman, Shea, Jette, Tassoni, Ottenbacher, & Fiedler, 1996) the MS Quality of Life-54 questionnaire (MSQOL-54; Vickrey, Hays, Harooni, Myers, & Ellison, 1995) and the Centre for Epidemiologic Studies Depression Scale (CES-D; Radloff, 1977). Although physical improvements concurrent with the resistance exercise undertaken by participants were observed, there was no effect seen in the EDSS, FIM, MSQOL-54 or CES-D. However, despite the data revealing a lack of support for the exercise and QoL relationship for those diagnosed with MS in this study, the authors stated that other types of exercise may improve QoL, and in fact referred to a previous study that found a positive relationship between aerobic exercise, MS and QoL (Petajan, Gapmaier, White, Spencer, Mjno, & Hicks, 1996).

Petajan et al. (1996) randomly assigned 54 patients diagnosed with MS to either a 15-week aerobic exercise or non-exercise group. Training consisted of 3 x 40-minute sessions per week of combined arm and leg ergometry and pre-and post-intervention assessments of fitness including maximal aerobic capacity (VO_{2max}), isometric strength, body composition, and blood lipids were measured. Additional measures included the Profile of Mood States (POMS) (McNair, Lorr, & Droppleman, 1992), Sickness Impact Profile (SIP; Bergner, Bobbitt, Kressel, Pollard, Gilson, & Morris, 1976), Fatigue Severity Scale (FSS; Chalder et al., 1993), Expanded Disability Status Scale (EDSS; Kurtzke, 1983) and neurological examination. The results of the investigation revealed that aside from improved bowel and bladder function in the exercise group, EDSS scores remained unchanged. There were no observed changes for both the exercise and the non-exercise group on the FSS. However, compared to baseline, the exercise group demonstrated significant increases in VO_{2max} ($p < .05$), upper and lower extremity strength ($p < .05$), and significant decreases in skinfolds ($p < .05$), triglyceride ($p < .05$), and very-low-density lipoprotein (VLDL; $p < .05$). For the exercise group, POMS depression and anger scores were significantly reduced at Weeks 5 and 10, and fatigue was reduced at Week 10. The exercise group improved significantly on all components of the physical dimension of the SIP and showed significant improvements for social interaction, emotional behaviour, home management, total SIP score, and recreation and past times.

One of the strengths of the study conducted by Petajan et al. (1996) was the range of physio-psychological measures undertaken. The interdisciplinary and comprehensive nature of the study design meant that this single investigation furthered knowledge regarding the magnitude of the impact of structured exercise upon a number of health

parameters for those diagnosed with MS. Fifteen weeks of aerobic training resulted in significant improvements across both psychological and physiological health parameters which the authors used to encourage physical activity as a feasible adjunct to traditional therapies for patients with MS.

However, the nature of the exercise undertaken by study participants may be difficult for people to adhere to without the level of supervision provided by this controlled study environment. Exercise participants engaged in three supervised training sessions per week for 15 weeks. Each training session consisted of a five-minute warm-up at 30% of VO_{2max}, 30 minutes at 60% of VO_{2max}, followed by a five-minute cool-down. Participants then performed five to ten minutes of stretching that concentrated on the posterior muscles of the lower leg, thigh and back. The average training intensity was 73% of maximum heart rate (HR_{max}) after ten minutes of exercise and reached 82% of HR_{max} towards the end of the exercise session. Although three weekly exercise sessions for exercise referral participants is feasible, it may be that reaching 82% of HR_{max} may be less achievable for some participants especially in the long-term and beyond the 15-week duration of the Patajan (1996) study.

Rather than an assessment of a specific relationship between a given chronic illness and QoL, Bowen et al. (2006) conducted a RCT exercise study with sedentary postmenopausal middle aged women (N = 173). The study utilised a combined programme of home and facility-based exercise and consisted of at least 45 minutes of moderate intensity aerobic exercise five days weekly for 12 months. Frequency and intensity of the exercise undertaken was monitored throughout the study and a variety of adherence strategies were used including: a) individual attention during classes, b) individual and group exercise behaviour change education classes, c) weekly phone calls, d) individual meetings at baseline and every 3 months to outline goals and provide feedback on progress, e) incentives such as water bottles, f) quarterly newsletters and g) tri-annual activities such as hikes. Women were randomly assigned to either the exercise group (n = 87) or a stretching only control group (n = 86).

Using the SF-36 Health Survey (Brazier et al., 1992) to monitor aspects of QoL, findings indicated that participants' perceived quality of life scores significantly

improved in several areas compared to control participants. These improvements were reported in the areas of mental health, general health perceptions and physical functioning. The changes were most pronounced at the three-month follow-up, although they persisted at the 12-month follow-up for general health perceptions and to a lesser extent for physical and mental health values. It is plausible that the decreases in physical and mental QoL perceptions in the longer-term could be accounted for by the consideration that participants will have become more familiar with their exercise behaviour and as a consequence, while perceptions of maintained general health continued, those perceptions of physical and mental health became less noticeable. In practice, it is commonplace for participants to report the greatest mental and physical gains during the early stages of referral although it is unknown if decreases in physiological response or perceptions of perceived benefit to exercise are responsible for the levels of exercise cessation reported within referral settings (e.g., Robison & Rogers, 1994).

Investigations of the role of exercise and QoL have been extended beyond middle to older age. Stathi, Fox and McKenna (2002) offered contributions to the literature in terms of the ways in which PA can make a difference to older adults through the implementation of a qualitative research design. Stathi et al. (2002) asserted that previous research located in the quantitative paradigm has revealed very little, and they proposed that these limitations often occur because the quantitative paradigm does not address an individual subjective knowledge about lived experience which is often revealed by personal accounts.

Stathi et al. (2002) recruited 28 community-dwelling, retired older adults, ranging in age from 62 to 81 years (M = 71 years; SD ± 5.35) all of whom attended individual and group interviews. The respondents included 15 women and 13 men who participated in an organised form of regular physical activity at least once per week. Cross-case analysis revealed 17 main themes with the key dimensions being identified as developmental, material, physical, mental, and social well-being. This study demonstrates how a qualitative approach was used to identify the complex inter-relations of subjective well-being and PA for older people. Physical activity appeared to contribute to the mental health of older adults through maintenance of a busy and active life, mental alertness, positive attitude towards life and avoidance of stress, negative function, and isolation. One of the strengths of the study design was to adopt a qualitative approach as this generated richness and depth of data and thus

adds to our understanding of the relationship between PA and well-being for older people. From a feasibility perspective, qualitative methods also negate some of the difficulties that often arise from alternative approaches that may require, for example, regular appointments with a medical professional for measurement purposes and the requirement of completing lengthy questionnaires.

In contrast, Chin A Paw, Van Poppel, Twisk, and Van Mechelen (2004) implemented a quantitative research design to examine the effect of different training protocols on quality of life, vitality and depression of older adults. Participants residing in the north-west Netherlands in long-term care facilities (N = 173, aged 64 to 94 years; M = 71 years; SD ± 13.44) were randomised either to six months of three different moderate-intensity group exercise training protocols, or to an 'educational' control condition. Exercise training consisted of two 45–60-minute training sessions per week of: 1) strength training, 2) all-round, functional training or 3) a combination of both. Perceived health (obtained by individual interviews), the Geriatric Depression Scale (GDS; Yesavage, 1982), the Vitality Plus Scale (VPS; Myers et al., 1999) and the Dementia Quality of Life questionnaire (DQoL; Brod, Stewart, Sands, & Walton, 1999) were administered at baseline and after six months.

After adjusting the score at baseline for age, sex and class attendance a significant difference between the strength training and the control group was found in the DQoL aesthetics sub-score ($p < .05$) The combined training group declined significantly in perceived health ($p < .05$), DQoL sub-score ($p < .05$), aesthetics sub-score ($p < .05$), esteem sub-score ($p < .05$) and VPS score ($p < .05$) compared to the control group. No significant differences between the all-round, functional training group and controls were found. Post-hoc analyses were conducted to examine the effect of the interventions including only those participants who attended at least 75% of the exercise classes (n = 74, i.e., 43%). After adjustment for baseline score, age and sex, no differences between the exercise groups compared to the control group were observed.

The results of this investigation suggest that neither strength training nor all-round, functional training of moderate intensity is effective in improving psychological quality of life, vitality or depression of older people living in long-term care facilities although there are a number of limitations to the study, some of which are recognised

by the authors. Chin A Paw et al. (2004) acknowledged that the multilevel analysis technique used to interpret the data may be useful for analysing discrete variables but that because ordinal outcome variables were employed, the results have to be interpreted with caution. The authors also identified the inherent difficulties in standardising and measuring the intensity of training and proposed that the lack of effect observed could be due to insufficient compliance to the programme by participants or insufficient intensity or frequency of the exercise itself. However, the authors did not fully explore the reasons for eight study participants discontinuing the intervention because they found that the exercise program was 'too intensive'. Additionally, for those participants who were included in the analyses, from the exercise logs it appeared that compliance to the strength training appeared difficult and participants subsequently often trained at a lower intensity and increased the intensity at a lower rate than prescribed - there was a mis-match between the strength training given to participants and what was actually undertaken over the course of the study. This highlights the importance of matching the appropriate mode, frequency and intensity of exercise particularly for clinical populations. In particular, it has been previously reported that feeling states are significantly worse at a higher intensity for the less active (Parfitt, Markland, & Holmes, 1994).

While GPs are becoming more familiar with the notion of exercise as a valuable prescription for disease prevention and management, the evaluation of the complete impact of prescribed exercise for clinical populations is unclear. This is perhaps particularly relevant for QoL outcomes given that generic measures of life-quality are unable to detect those constructs that are of particular relevance to those who have been referred for exercise for health improvement. The complexity of assessing life-quality via the use of questionnaires in terms of measurement and interpretation of the data are not limited to that which is described in the Chin A Paw et al.'s (2004) research. Other than appropriate levels of reliability and validity, selecting a measure to use within applied and/or research settings can often be the result of the measures available for the target population and practitioner/researcher personal preference. The QoL measurement challenges experienced by exercise referral professionals are discussed in the following section.

1.8 Health as a Holistic Construct: Matching the Tools to the Task

According to the NQAF (DH, 2002), alongside other services, exercise referral schemes will make an important contribution to improving the nation's health. The most commonly referred to and widely accepted definition of 'health' is provided by the World Health Organisation (WHO) which states that health is "a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity" (WHO, 1948, p. 2). This definition has not been amended since 1948 and from an applied perspective is generally regarded as the most relevant and inclusive for health practitioners, and is consistent with the bio-psycho-social paradigm of illness in medicine (Engel, 1977, 1990; Hojat, Samuel, & Thompson, 1995). However, subsequent publications have proposed that the WHO (1948) definition of health may more accurately describe perceptions of happiness (Saracci, 1997) and more recently, that attempts to define health may be a futile exercise and may not be able to capture health's true complexity (Alejandro, 2008).

One of the key challenges that is evident from a review of the literature is the variance of perspectives regarding the notion of health. In the same manner that Hunt (1997) has asserted with respect to defining life-quality, when considering if an intervention has an impact upon a given health parameter, we must first have a clear understanding of the definition of that parameter. As described previously, the efficacy of an exercise referral scheme is often judged by evidence of participant physiological change, assessed via the collection of prescriptive physiological measures. In this instance, a referral for exercise can only be considered as having the capacity to impact upon the physical element of 'health' as defined by the WHO.

The importance of physiological outcome measures has been described by Winett (1998) who stated that scientifically based behavioural prescriptions if simply adhered to "would yield the desired outcome, that is, enhanced fitness and alteration of body composition" (Winett, 1998, p. 209). The difficulty with this claim is twofold: firstly, the notion of adherence is not given full consideration by Winett, 1998) in terms of the role of the motivational and behavioural change processes that surround it, and secondly, Winett (1998) presumes that the "desired outcome" is physiologically universal and unchanging for each participant. Given the mental and physical characteristics of those referred, this is a viewpoint that restricts the

potential responses to and outcomes of exercise. What is more, it is suggested that it is incorrect to suppose that "simply performing *more* of an activity, typically lower to moderate intensity activity, is the way to improve health outcomes" (Winett, 1998, p.211) and that "what is particularly distressing about conclusions, based on the exercise determinants literature, is that most of the participants in these self-report survey studies were sedentary or did not exercise regularly. Having little or no experience with proper exercise, they were merely reporting on their perceptions" (Winett, 1998, p.214).

A broader understanding of health is likely to demonstrate a range of desired outcomes, including physical, mental, social and clinical factors. These desired outcomes should be identified by the client and the ability to achieve them should be supported by suitably skilled exercise professionals. Individually reported perceptions of the perceived benefits of a slowly developing active lifestyle for sedentary individuals should be encouraged to support self-efficacy and enhance motivation and adherence. This level of perception does not require any previous experience of exercise and not only reflects good practice, but the more holistic understanding of health as proposed by WHO (1948) and advocated by Chandra (2001). What these differing understandings of health – and in particular health outcomes demonstrate, is the subsequent difficulty with assessing the efficacy of an exercise intervention upon something that differs in its comprehension (i.e., health). This may account for the evidence base to support the efficacy of exercise referral schemes which is often regarded as 'conflicting' or 'inconclusive' (Crone, Johnston, & Grant, 2004; Graham, Dugdill, & Cable, 2005).

In addition to the differing notions of what constitutes an understanding of 'health', there are also frequent inconsistencies in language within the literature that make it a difficult task to understand how exercise referral schemes impact upon the diverse areas of health and to what degree they are effective in these areas. For example, Thurnston and Green (2004) proposed that research to date suggests that the evidence base for exercise referral schemes' impact is limited; however, what is difficult to establish is what is precisely meant by 'impact'. The term can encompass physical health changes, psychological improvements, PA and exercise behaviour change, long-term adherence, quality of life and environment, for example. The evidence base for each element that may fall under the category of 'impact' for participants is varied. This makes it difficult to untangle and establish with any

degree of certainty what exactly the authors are claiming exercise referral schemes have a limited impact upon.

Similarly, Harrison, McNair, and Dugdill (2005) ascertain that experimental evidence questions the effectiveness of exercise referral schemes. In this instance, it is unclear as to what health parameters exercise referral schemes have the most and least effect upon. The term 'efficacy' itself is typically used in a very generic context and if we are to further our understanding as to how to help people change and maintain exercise behaviour a greater amount of clarity regarding the exact processes and or health outcomes being investigated is required. In the interests of communicating the outcomes of scientific investigation transparently and with a high degree of clarity, exercise referral schemes should be assessed in terms of specific and identifiable parameters. For example, physical, emotional, mental, social, environmental, economic, and behavioural factors – a holistic approach. The necessity for clear reports of evaluation methods and transparency has previously been noted (e.g., Breckon, Halley Johnston and Hutchison, 2008; Jolly, Duda, Daley, Eves, Mutrie, Ntoumanis et al., 2009). Greater clarity regarding the specific focus of the health parameters under investigation alongside the methods used would help to identify what the strengths and weaknesses of a referral into exercise are for those who have experienced it, and reduce the confusion over generic references to 'efficacy', 'impact' and 'effectiveness'. Furthermore, employing the most appropriate data collection tools with which to evaluate these parameters is also paramount.

The methodological challenges experienced in disseminating evidence-based interventions to promote PA have been identified previously (Rabin, Brownson, Kerner, Russel, & Glasgow, 2006). It is proposed that the exercise-specific QoL measure developed as a result of the current research will contribute significantly to standardising yet broadening the way in which exercise schemes are delivered and evaluated, and enable practitioners to have the right tools for the task at hand. In addition, academic communications made using data generated from the measure may limit the confusion regarding the efficacy, impact and effectiveness of such schemes as the instrument will provide some continuity and specificity with respect to psychological assessment. In addition, adopting an approach to evaluation that utilises the more common physiological assessments alongside the psychological tool

developed for the current research will not only represent the more holistic notion of health but may also be suitable for the purpose of internal audit and to secure future funding.

1.9 Summary of the Literature Review

Published work has reinforced the critical necessity to address the current shortfall in the consistent and routine assessment of psychological outcomes for referral participants and has provided insights into the history of the development of exercise referral schemes, variance in evaluation methods and the strength of the relationship between exercise and life-quality, particularly for clinical populations, while balancing this relationship with the recognition of challenges in defining and measuring the concept.

In particular, a critique of the exercise referral evaluation literature has revealed important findings. Practitioners have been encouraged to address the psychological processes and outputs of a referral into exercise and guidelines have been published (DH, 2002). However, these guidelines are considered limited which may account for practitioners feeling unsure of the best measures to use when assessing effectiveness (Dugdill et al., 2005). There have been increasing calls for broader, more holistic evaluations of exercise referral schemes (Dugdill et al., 2005; Oakley, 2001; Springnett & Dugdill, 1995; Taylor, 2003) and one of the most common responses in practice has been to introduce generic measures of life quality (e.g., SF-36; Brazier et al., 1992; Ware, 1996) and the EuroQoL (EQ-5D; Brooks, 1996) into evaluation protocols.

The literature revealed two key challenges regarding QoL assessment - namely, the definition (e.g., Cohen et al., 1996), and the measurement of the concept (e.g., Allison et al., 1997; Holmes, 2005). However, despite the continued debate regarding these challenges, the necessity to capture indicators of mental and physical well-being within healthcare settings is such that it is sufficient to take precedence over these philosophically focused debates (Allison et al.,1997). Indeed, in the context of exercise and health improvement the requirement to have QoL measurement tools that are tailored and sensitive to the population for which they are intended is

reflected in the continuing development of such tools (e.g., Doward et al., 2003; Hilditch et al., 1996; Kolotkin et al., 2001; Las Hayas et al., 2006).

Additionally, the literature demonstrates a sound evidence base for the link between PA, exercise and QoL, particularly for clinical populations and those with health conditions that are typical of those who are referred to a UK referral scheme (e.g., Milne, Guilfoyle, Gordon, Wallman, & Courneya, 2007; Motl & Snook, 2008). Furthermore, measures of life-quality have been deemed as the primary end-point in clinical settings (Pais-Ribeiro, 2004). While GPs are becoming more familiar with the notion of exercise as a valuable prescription for disease prevention and management, the evaluation of the complete impact of prescribed exercise for clinical populations is inconsistent and therefore unclear. This is perhaps particularly relevant for QoL outcomes given that generic measures of life-quality are unable to detect those constructs that are of particular importance to those who have been referred for exercise for health improvement.

The following chapter provides an overview of the specific procedures undertaken to develop the new scale and the initial phase is also reported in detail. The first study utilised a qualitative focus group approach to generate data that was used to develop items for the test measure. These focus groups were attended by participants who had completed their initial 12-week referral for exercise at an exercise referral scheme typical of those found in the UK.

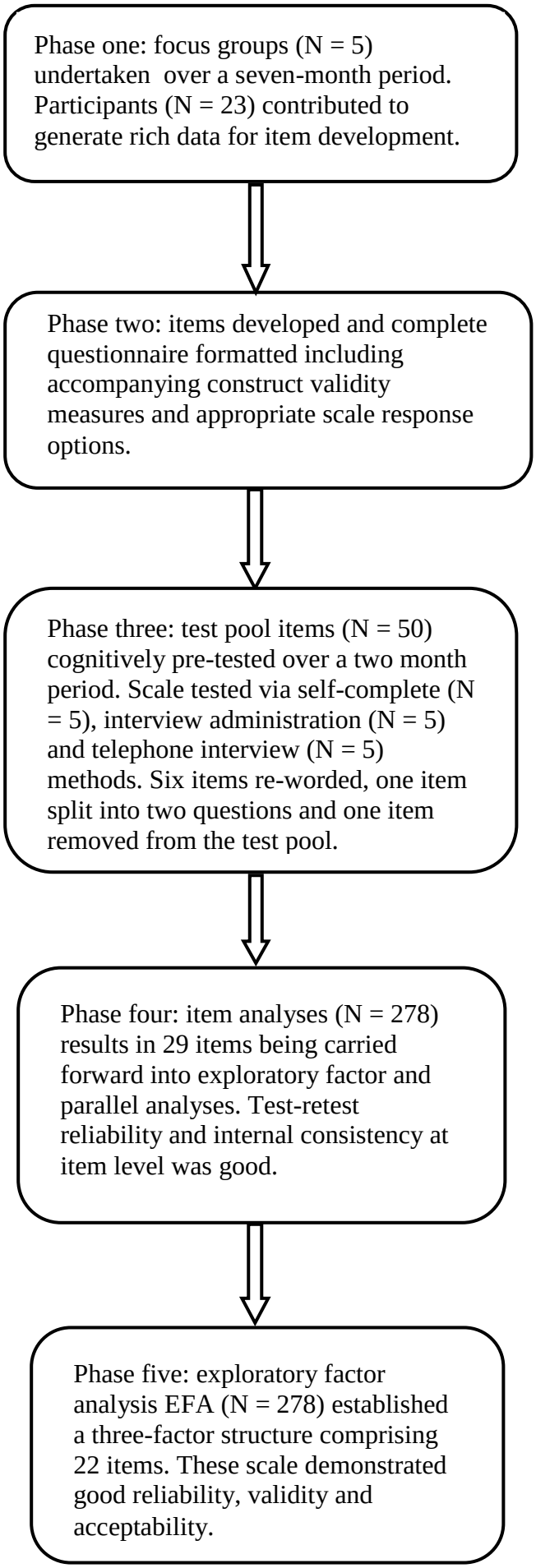
CHAPTER II Scale Development:

Overview and Focus Groups

2.1 Scale Development: Overview

The following four chapters report on the specific procedures undertaken to develop the new scale. Figure 1 provides an overview of these distinct phases of research which were informed by best practice guidance for new health measurement scale development (Streiner, & Norman, 2008).

Figure 1. Procedures undertaken to develop and validate the final scale



2.1.1 Introduction: Scale Development – Focus Groups

The primary purpose of phase one of the current research was to generate items (questions) that would comprise the initial test pool for the exercise referral sensitive QoL measure under development. Miles and Huberman (1984) have asserted that having an understanding of what it is a researcher wants to find out inexorably leads to the question of *how* you will get that information. Miles and Huberman's (1984) prior structuring strategy of a research design categorises the focus and purpose of the current phase of research as an open question or exploratory study and as such this type of study is required to be far less structured than confirmatory studies (Miles and Huberman, 1984; Silverman, 2000).

In contrast to a quantitative study that may ask 'why' and look for controlled group comparisons, a qualitative study usually indicates a research question of 'how' or 'what' (Creswell, 1998). Creswell (1998) also recommends a qualitative approach if the topic is one that requires exploration and if there is a necessity to present a detailed view of the topic. The focus of the current phase of research was to generate a detailed exploration of the common responses to and outcomes of exercise referral with regards to perceptions of life-quality - *how* and *in what way* does exercise referral contribute to perceptions of exercise-related life quality? Therefore, a qualitative approach was most fitting.

It seems logical that if self-report data regarding a given research topic is required, that one need only ask those who are best placed to comment. In this case, those who had completed a 12-week exercise referral programme. Denzin and Lincoln (1994) suggested that this purposeful sampling method of recruiting participants is typical of qualitative approaches to research whereby groups, individuals and settings are sought to represent the processes being studied and where they are most likely to occur. When self-report qualitative data such as this is required an individual or group interview (focus group) approach is usually employed.

Maintaining the focus of conversation while adopting a flexible approach to conversation and building good rapport requires experience and skill. This skill is central to both the individual and focus group method of data collection. Both methods have the potential to generate the richest data, although poor interviewing

skills also run the risk of generating too many data that is not useful to the research question (Holloway, 1997).

To facilitate rapport and trust building interviewers may provide some personal context to the nature of the research and how they came to be involved in the research process and this may also include describing any personal connections with the topic under study (Grbich, 1999; Thompson, 1995). The intimacy and trust associated with qualitative interviewing is the opposite to the distance and control valued in quantitative research (Corbin & Morse, 2003) and this approach has produced many important findings for health-related research (e.g., Mischler 1986; Morse, 1989).

It is commonplace for a focus group approach to be employed for the purposes of generating items for a questionnaire under development (e.g., Las Hayas et al., 2006; McKenna, Doward, Whalley, Tennant, Emery, & Veale, 2004; Spies, Coyne, Guaou, Boyle, Skyrnarz-Murphy, & Gonzalves, 2002). The group interaction associated with focus groups provides a number of advantages: conversations between participants may generate new discussions that contribute to understanding the research question (Hansen, 2006), listening and engaging in conversation may assist participants to clarify their point of view (Bates, Lynch, Bevan & Condit, 2004; Denning & Vershelden, 1993) or comment on sensitive topics (Kitzinger, 1994) and focus groups have the ability to produce more data in the same time than would an individual interview (Holloway, 1997). For these reasons, a focus group approach was chosen in favour of the individual interview for the current research.

The following sections provide a critique of a focus group approach to data collection, reflexivity, data analysis and representation and rigour. This critique helped to inform the methods employed for the current study.

2.1.2 The Focus Group Approach

Many attempts have been made to define focus groups. However, a common theme is that they are small groups of people who possess certain characteristics who meet to provide qualitative data in the context of a focussed discussion (Krueger, 1994). They are group interviews. Typically a group facilitator or moderator guides the group discussion to explore a particular research question (Morgan, 1997). The moderator is responsible for creating a supportive environment that encourages participants to share their opinions and experiences (Basch, 1987; Kitzinger, 1994). The focussed discussion generates data that is usually audio-taped, transcribed and subject to content or thematic analysis. As mentioned previously, one of the central differences of focus groups to individual interviews is the interaction between participants (Rice & Ezzy, 1999) and it has been suggested that this may assist the participants to clarify and explore their views (Kitzinger, 2000). However, participants that are too dominant or disengaged from discussions can be problematic and transcription of the focus groups can be challenging in terms of identifying which of the participants is speaking and also the volume of data that is produced (Hansen, 2006; Holloway, 1997).

In practice, recently focus groups have been included into routine evaluations of public health interventions as practitioners increasingly recognise the importance of involving users in the development and evaluation of health services.

2.1.3 Data Collection

Interview guides or questioning routes (Krueger & Casey, 2000) are often produced pre-interview to facilitate the data collection process. These are far more flexible and dynamic than questionnaires but still require careful planning and consideration (Loftland & Loftland, 1984). In qualitative interviewing, good questions are often open or open-ended, meaning that the question discourages respondents from providing a short answer. They are also never leading or overly directive (Hansen, 2006). It is typical that an interview begins with broad descriptive questions and follows with probing questions to elicit further detail (Krueger & Casey, 2000; Loftland & Loftland, 1984). Probes are intended to prompt elaboration, to direct the conversation into an area that has received comparatively less discussion or to

encourage the respondent to discuss a new topic area. Probes may be included in an interview guide so as to remind the interviewer to follow-up on anticipated discussion points and others may be introduced ad hoc as the interview progresses (Hansen, 2006).

The manner in which qualitative data is collected is of critical importance. Considerations such as the level of rapport between the interviewer and those being interviewed, how questions are delivered etc. contribute to the quality of data generated. In this respect, the integration of psychological counselling skills into interview settings has been given some attention (e.g., Hansen, 2006). As it is commonplace for health professionals to have received such training (Hansen, 2006) and many of the techniques employed by person-centred styles of counselling in particular, mirror what is deemed as good interviewing practice.

Data recording is central to the method of qualitative research. Data needs to be recorded in some way so that the researcher can review it retrospectively and so that findings are accountable and transparent (Hansen, 2006). Interviews, perhaps more especially focus group interviews are usually audio recorded (Hansen, 2006; Holloway, 1997) and most recently, qualitative researchers have started to make use of digital recording (Hansen, 2006). Digital recordings tend to have much better sound clarity (Stockdale, 2002) they are more flexible in terms of recording times and practically they are easier to transport to and from the interview venue (Hansen, 2006). Mitchell, Peterson and Kaya (2004) and Muhr (2000) have also suggested that digital recording may also make transcription easier. Data that may be vital to the research area under investigation may be lost without audio recording including the narrative itself, intonation, nuance, meaning and sequence.

On occasion interviews may also be video recorded so as to provide some visual context to the topics being discussed (Holloway, 1997). Both the individual interview and the focus group approach tend to adopt a ‘data saturation’ approach to establishing the number of interviews that are required to represent the research area adequately. Researchers often believe that data saturation is achieved when a concept is mentioned frequently, described in similar ways by different people or when the same ideas arise repeatedly (Holloway, 1997). There is no method of predicting when data saturation will occur (Holloway, 1997). However Calder (1977) suggested that within the context of focus groups, it is typical for data saturation to occur by the

third or fourth group. For the current study, data saturation was gauged in the manner identified by Holloway (1997) – when similar concepts were mentioned frequently across the focus groups and related responses were provided to the questions asked.

Once the data has been recorded, it is typical that the recordings are transcribed and it has been suggested that the 'gold standard' is a full transcription (Hansen, 2006) also often referred to as verbatim transcription. On occasion, the transcription may also include conversational pauses and filler sounds (Lynch, 1993) to facilitate the full context of the dialogue and assist with interpretation. The transcription process itself is considered by some researchers as being part of the analytical process (Riessman, 1993) and helps to familiarise the researcher with the data.

Qualitative researchers often use the term 'generating' data in preference of data collection (Mason, 1996). This is perhaps because qualitative researchers do not merely collect and describe data in a neutral and detached manner but rather, the researcher is considered as actively constructing knowledge (Mason, 1996). Arguably, the most important data collection tool in interview method studies is the interviewer and qualitative researchers often reflect upon their own attributes, personal characteristics, knowledge and life-experiences in terms of how they may impact upon the interviews they conduct. This self-awareness is often referred to as reflexivity and is often considered a valued part of the qualitative research process.

2.1.4 Reflexivity

Researchers are reflexive when they critically examine their own assumptions and actions regarding the research process (Holloway, 1997). Reflexivity is often considered as an essential part of qualitative research as the researcher is the main tool in the data gathering process. He or she uses 'the self' as an instrument (Holloway, 1997). Indeed, the role of the researcher may be subject to the same critical analysis as the research itself (Rice & Ezzy, 1999). Reflexivity encourages honest consideration about the researcher's role in their project (Shacklock & Smyth, 1998) and it is also recommended that reflexive accounts are written in the first person, using the pronoun 'I' (Horsburgh, 2003; Webb, 1992) as in this way a researcher may assert their 'ownership of responsibility' for their views (Berg, 2004). Reflexive accounts are often published separately to research findings, acting as worthy contributions to scientific understanding in their own right (e.g., Carolan, 2003; Hand, 2003; Haynes, 1999).

2.1.5 Approaches to Data Analysis and Representation

There is a considerable amount of variation in the different approaches employed by researchers to analyse qualitative data and there are no standardised rules as to how qualitative data should be analysed (Hansen, 2006). Rather, the researcher may be most concerned with assessing how best to represent the data. It may be particularly helpful to ask 'what is the purpose and function of the data?' to assist with choosing an appropriate method of analysis. Four broad approaches to analysis include: content analysis, iterative/thematic analysis, narrative analysis and discourse analysis, however each of these may include sub-methods. For example, grounded theory is an example of the iterative/thematic approach to data analysis (Hansen, 2006).

The central feature of frequency measurement is more commonly associated with quantitative approaches to analysis (Grbich, 1999). Inductive content analysis involves the derivation of themes and constructs from data without imposing a prior theoretical framework or without counting (Holloway, 1997). The term deductive content analysis is often used to refer to content analysis that employs an *a priori* theoretical framework. However, since approaches such as grounded theory and

phenomenological approaches have been more fully developed inductive or deductive content analysis is usually referred to in these terms (Holloway, 1997).

Iterative/thematic analysis involves identifying ‘themes’ in data. Open coding is the first stage of this approach to analysis (Hansen, 2006; Holloway, 1997). Closely linked concepts are grouped into categories and following this stage, the categories are collapsed into themes or constructs (Holloway, 1997).

Narrative and discourse approaches to qualitative analysis are just as concerned with how something was said and the reasons why it may have been said in such a way as with what was said. Such analyses may also take into consideration the social and political realm in which the data were collected and help to inform cultural and individual ideas about the social climate in this respect. Indeed, Hansen (2006) accounts for the growing interest in the narrative or discursive approach to analysis for these reasons.

The development of computer software such as NVivo, for example have been designed to assist qualitative researchers to record, store, index, code and sort large volumes of data (Schwandt, 2001). Regardless of the analytical approach adopted to making sense of the data, software for qualitative analysis can benefit the researcher in terms of speed, consistency, rigour, and access to analytical methods that are not available by hand (Weitzman, 1999).

2.1.6 Rigorous Qualitative Research

The application of quantitative understandings of reliability and validity are inappropriate for the evaluation of qualitative research as the purpose and focus of the paradigms are not directly comparable (Horsburgh, 2003). Furthermore, Holloway (1997) has asserted that reliability in terms of consistency is difficult to achieve in qualitative research because the researcher is the primary research instrument. That said, the notion of reliability (i.e., repeatability) alludes to the scope for the transferability of the study from one specific case to another. In this respect, Lincoln and Guba (1985) maintain that this may be achieved if the researcher provides adequate detail about the circumstances of the situation or case that was studied. Of course, although the word limit of scientific journals for qualitative

reports is often higher; this in itself can create difficulties for transferability in the way that Lincoln and Guba (1995) suggest. In contrast to the notion of some level of transferability though, Popay, Rogers and Williams (1998) have argued that the hallmark of good qualitative research is its variability rather than its standardization.

Purposive sampling may help to improve the rigour of a qualitative investigation as Rice and Ezzy (1999) have previously commented that qualitative research is not concerned with producing findings that can be statistically generalised to the whole population. Equally, transparent reporting of methods and analysis is important for the rigour of qualitative studies (Hansen, 2006). For focus group and individual interview studies, this includes sufficient use of participant quotations to exemplify data interpretation. Such thorough and transparent reporting of findings demonstrates the intellectual integrity of the research and lends considerable credibility to the final set of findings (Patton, 1990). Additionally, the use of reflexive accounts (Webb, 1992, 1996) and audit trails (Koch, 1994) may also contribute to the rigour of qualitative research.

2.1.7 Summary of the Introduction

The exploratory nature of the current phase of research indicated that a qualitative approach was best suited to generating a volume of participant perspectives regarding the how and why of exercise referral and perceived life-quality (Miles and Huberman, 1984). The literature has emphasised the benefits of employing a focus group approach in terms of the added value of participant interaction (e.g., Bates et al., 2004; Denning & Vershelden, 1993; Hansen, 2006; Holloway, 1997; Kitzinger, 1994), their ability to assist participants to clarify their point of view (Bates et al., 2004), comment on sensitive topics (Kitzinger, 1994) and to produce more data in the same time than would an individual interview (Holloway, 1997).

Consequently, this approach is the most suitable for the purposes of open exploration and the production of rich data for the purposes of item generation. More specifically, an iterative/thematic approach to analysis employing an open coding approach initially is likely to yield the richest and most appropriate volume of data for the purposes of item analysis (Hansen, 2006; Holloway, 1997).

The literature has also indicated the importance of the role of the researcher in the enquiry process, not only in terms of interviewing skill, but also in terms of a commitment to a reflexive approach. Allowing the researcher to critique the impact of the researcher upon the research process and subsequent analysis and interpretation. This approach to the research may also help with contributing to the rigour of qualitative research (Webb, 1992, 1996) in addition to purposive sampling methods (Rice & Ezzy, 1999) and audit trails (Koch, 1994) and transparent reporting of methods and analysis (Hansen, 2006). These considerations informed the methods employed to the current study.

2.2 Aims

The aims of this study were to explore the common responses to and outcomes of a referral into exercise with respect to perceived exercise-related life-quality. The end point aim of this qualitative phase of research was to generate items for the development of an exercise referral sensitive life-quality scale.

2.3 Methods

Ethical approval was obtained from Nottingham Trent University and the UK National Research Ethics Service (NRES; reference number 06/Q2404/36; date of approval 25/4/2006). Five focus groups were conducted in total. Participants (N = 23) for all five focus groups were recruited through a local authority exercise referral scheme in Nottingham, UK. These participants did not consist of an existing group of exercisers, but rather were recruited individually and were selected to attend a group based upon their availability. An iterative process of gauging data saturation as described by Holloway (1997) informed when enough focus groups had been conducted to fulfil the aims of the research. All participants were provided with an NRES template information sheet describing the study (Appendix 1) and also a shortened summary version (Appendix 2). All participants provided full written consent to their participation, again using the NRES consent form template (Appendix 3).

In the interests of producing the richest data, a focus group approach that utilised many of the collaborative and exploratory principles of an unstructured in-depth interview was employed. Based on the recommendations of Krueger and Casey (2000) a questioning route was developed. The questions were designed to guide the focus group participants through the interview gradually via structuring the enquiry around introductory, transition, key and ending questions. The primary investigator formulated questions that were designed to reflect these categories and these were then subject to a process of peer de-briefing (Spall & Stephen, 1998). This was an iterative process whereby members of the research team, experienced in the methodology of using focus groups as a means of scale construction provided comments and suggestions for amendments to the questioning route until the final version was produced. This can be reviewed below.

Following a welcome and the introduction of each group member by themselves the primary investigator will ask the following questions:

Introductory Questions:

1) What does physical activity mean to you?

2) I am looking to explore peoples' responses to, and experiences of exercise referral. Firstly, I would like to know what your views are about responses to and experiences of exercise referral?

Note to Primary Investigator: Be aware of responses indicating expectations, responses to exercise and others.

3) I want you all to recall your first experience of your referral into exercise, perhaps your first appointment with your instructor. Can you describe or explain what you were thinking and feeling?

Note to primary investigator: Probe for perceptions of possible expectations.

Transition Questions

1) You have provided some really useful information as to your thoughts and feelings at the early stages of your referral (provide examples). I was wondering how these thoughts and feelings may have changed over the course of your programme, if at all?

Key Questions:

1) I'm going to ask each of you the same question in turn now. Can I ask what the term quality of life means to you? There is no correct answer; I am really interested in what each of your opinions are.

Note to Primary Investigator: Ask each participant 'what does the term quality of life mean to you?'

2) Before starting your programme of exercise, what was important to you?

3) I want you to consider your health as taking into consideration your social, emotional, physical and daily living needs, in addition to any illness you may have. I'm wondering how your referral into exercise may have impacted on these things if at all?

If we can begin with your social health?
Emotional health?
Physical health?
Daily Living?

Note to primary investigator: Emphasise text in bold and speak slowly when describing each need. Address each need in turn. Probe for physical, emotional and active daily living (ADL) outcomes. Offer encouragement for contributions.

4) So to summarise, in your opinion, how would you list the main outcomes of your exercise referral? If you prefer, you can list 1-10 for example.

Note to primary investigator: Encourage participants to list the most salient outcomes (i.e., 1-10).

5) Remember earlier we discussed what was important to you before starting your programme of exercise? Considering that you have completed your initial exercise referral period, what's important to you now?

Ending Questions:

1) Suppose you had one minute or less to describe your experience of exercise referral, what would you say?

Note to Primary Investigator: Allow each participant approximately 1-minute to respond.

2) Of all the things we have discussed today, which is most important to you?

3) Note to primary investigator: Provide a 2-3 minute oral summary of issues evoked by the key questions then ask:

3a) How well do you feel what I've said captures what has been reported by you today?

4) Note to primary investigator: provide an overview of the purposes of the research then ask:

4a) In light of this, is there anything we should have talked about but didn't?

Note to primary investigator: Delivery is critical, allow the participants time for reflection.

Thank participants and informally ask how they feel about their participation in the group.

Based upon the recommendations of (Grbich, 1999; Thompson, 1995) to further enhance rapport, prior to the start of the interview the primary investigator provided some background to the development of the research including personal accounts of her history in designing the exercise referral scheme that the participants had attended and how difficulties with psychological and holistic evaluation had resulted in her registering for a PhD. Participants were also reminded that any information they provided during the interview would not impact upon changes to funding the scheme or in any way that would impact negatively to the scheme development. These accounts were not written into the questioning route so that each introductory explanation was communicated naturally in a personable manner and any temptation to read such content from a script was avoided.

Throughout the focus groups, questioning probes and invitations for elaboration were introduced by the primary researcher. The primary investigator has previously employed a focus group approach for the purposes of evaluative research and was engaged in peer supervision regarding focus group conduct throughout the period of the current research. Peer supervision and ongoing reflection/modification of the primary researcher's focus group skills contributed to the techniques and conduct employed in the current study. Additionally, the primary investigator was trained in various counselling styles, primarily from a person-centred perspective and so some of these techniques were used to help facilitate the early development of rapport between the facilitator and group members, ensure questions remained open, probes were introduced appropriately alongside opportunities for participant reflection and to add to the overall quality and richness of the data.

The five focus groups were all conducted on weekday evenings during the months of July to January. In each case, the focus groups were conducted in a quiet meeting room at a leisure centre centrally located to the area in which participants resided. A co-facilitator was not utilised for any of the focus groups as the largest number of participants attending any one group was six. This number of participants was considered manageable for a single facilitator.

Four female participants were recruited for an initial pilot focus group. Participants ranged in age from 43 to 74 years ($M = 62$, $SD \pm 14$). The pilot group was conducted to address the relevance and accuracy of the questioning route, establish the

emergence of initial themes and provide a valuable opportunity for reflection upon the interviewing process. Based upon the recommendations of Hansen (2006), the pilot phase was included as an important and integral part of the complete set of focus groups (N = 5) including the subsequent analysis as it is difficult to quarantine an interview or set of interviews in inductive research designs where all data can be seen as part of the overall study (Hansen, 2006).

Opportunity for reflection and an assessment of the primary researcher's role within the process was also established via the use of a diary. Entries were made in the diary both before and after focus group completion to allow the primary researcher to accurately and openly reflect upon her understanding of any difficulties encountered and lessons learned. For example, an entry reflecting upon the start of the first focus group that was conducted after the pilot group read: "looking forward to analysing the data. Feeling that my early intuitions of common reports will help with my line of questioning/probes in the next group." According to Webb (1992, 1996) this process enhances the rigour of qualitative research and Finlay (2002), has argued that this reflective process also addresses the integrity and trustworthiness of the research.

2.3.1 Participants

Participants were purposively sampled from a local authority exercise referral scheme. According to Rice and Ezzy (1999) this method of sampling contributes to the rigour of qualitative research because the demographic, and experiential qualities that participants bring to the research have been identified as directly relevant to the research question. The exercise referral instructor who was responsible for supporting attendees through their referral programme facilitated recruitment by acting as a gatekeeper between the participants and the primary investigator. With permission from each interested person, the primary investigator then contacted these participants, explained the purpose of the study and the requirements of their involvement in a focus group in more detail and those who agreed to participate thereafter were given a choice of dates to attend a focus group. This method of follow-up phone calls and flexible attendance options is an approach to successful recruitment that has been recognised previously (Hansen, 2006).

The data were transcribed to explore the emergence of initial themes and address any necessary adjustments to the questioning route. For the purposes of the primary research, male and female participants were identified in the same manner as those recruited for the pilot group. Individuals who completed 12 weeks of their exercise referral programme provided a continuous source of potential participants. Those who had completed their initial referral period of 12 weeks were selected as it was considered that these individuals would have had adequate experience of their referral into exercise to be able to comment on the scheme's impact upon exercise-related life-quality. Focus groups were conducted until the fifth group when responses began to elicit similar and common ideas to the point of 'data saturation', as typically occurs by the third or fourth group (Calder, 1977).

One pilot focus group and four subsequent groups were conducted, which included a total of 23 participants (F = 15; M = 8). Participants ranged in age from 35 to 77 years (M = 60, SD ± 10) which is representative of individuals who are referred into structured exercise. Health difficulties for which participants were referred for exercise were also typical of exercise referral populations (Table 2). The employment status of participants included employed (n = 7), retired (n = 7), voluntarily employed (n = 1), self-employed (n = 2), self-employed-retired (n = 1), unemployed (n = 2) and unknown (n = 3). Reasons for a referral into exercise ranged from joint difficulties, weight loss and post-operative rehabilitation to heart health and cardiac rehabilitation (Table 2).

Table 2. Demographic characteristics of exercise referral focus group participants (*N=23*).

Focus Group	Gender	Age	Reason for Referral
Pilot	Female	43	Fibromyalgia, Osteoarthritis
Pilot	Female	74	& Replacement Knee
Pilot	Female	61	Asthma
Pilot	Female	71	Weight Loss
1	Male	54	Post-Operative Rehabilitation & Fitness Gain
1	Male	52	Weight Loss & Hypertension
1	Female	70	Cardiac Rehabilitation
1	Female	58	Weight Loss
1	Female	77	Joint Stiffness
1	Female	65	Weight Loss
2	Female	53	Back/Hip Problems
2	Female	58	Arthritis, Post-Operative Preparation, Weight Loss & Under-Active Thyroid
2	Female	66	Increase Fitness & Physical Activity Level
2	Male	73	Knee Joints
2	Female	64	Weight Loss
3	Male	55	Heart Attack Rehabilitation
3	Female	49	Hypertension
3	Male	62	Heart Health & General Fitness
3	Male	57	Improve Fitness Following Illness
4	Female	60	Fibromyalgia
4	Female	35	Arthritis
4	Male	59	Post Surgery Rehabilitation
4	Male	67	Weight Loss & Sedentary

2.3.2 Data Handling and Analysis

The focus group data were transcribed verbatim by a transcription service external to the research team. However, to gain familiarity with the data and to check for errors, the interview scripts were read a minimum of three times and the audio material reviewed thoroughly. This familiarisation and error checking process is one that has been utilised previously when the transcripts are produced externally (e.g., Hansen, 2001).

An inductive iterative/thematic analysis approach was employed to the analysis of the transcripts and audio data. It was recognised that for the purposes of item generation, the depth and level of analysis required is not representative of, or comparable to stand-alone qualitative studies. For example, the exploration of the relationship between themes was not necessary. However, the analytical procedures undertaken were typical of those employed for qualitative data analysis. The analysis was conducted in two phases and the data were treated as five complete sets (one set per focus group). An inductive approach to analysis allowed the data to identify how exercise was related to perceived life-quality without the use of any underlying hypotheses, or for the purposes of theory development as would be indicative of a deductive approach to analysis (Gilgun, 2002).

The first phase employed an open coding level of analysis which served to produce a wealth of rich data. During this phase, with the use of NVivo, the data were coded according to what aspect of exercise-related life-quality participants were referring to during their conversations. For example, 'choice of activity' and 'improvements in confidence.' Where relevant, some dialogue was coded into more than one category when it was clear that participants were describing two or more related issues. For example, participants described having experience of being physically active in the past alongside perceptions of improved confidence regarding the use of exercise equipment. In such an instance, the passage of dialogue was coded as both 'history of physical activity' and 'improvements in confidence.' Such a coding process would normally allow for the exploration of related concepts for the purposes of theory development, for example. However, this was beyond the scope of the aims of the current study and the open coding level of analysis was sufficient for item generation. Throughout the open coding and subsequent higher order thematic phase

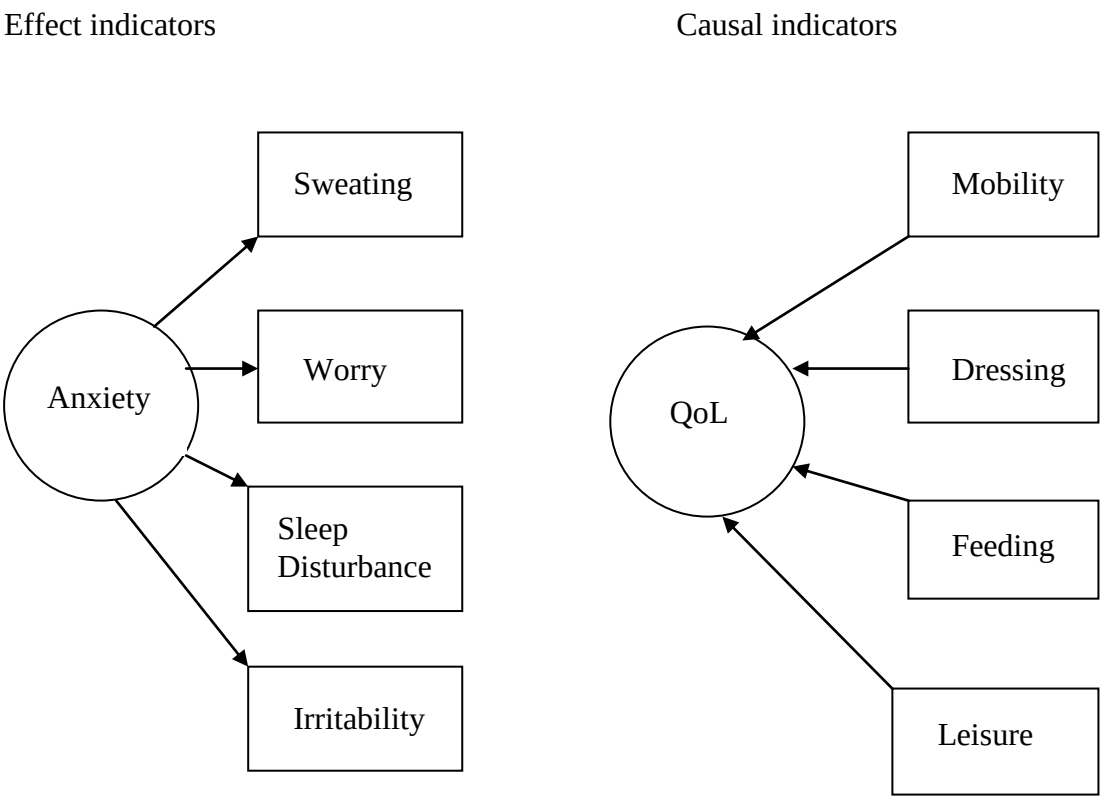
of analysis, the titles given to coded data underwent an iterative phase of peer de-briefing (Spall & Stephen, 1998). This allowed for the research team to collaborate and reach a consensus regarding the most appropriate title to reflect the content of the coded data and how one theme was decided to be distinct from another.

The second phase of analysis translated this open coding level of data into a smaller number of broad categories. In some instances this required collapsing categories that were clearly describing the same concept. For example, 'mode of activity' and 'choice of activity.' In other instances categories were combined to reflect a related yet broader concept which is typical of this level of analysis. This second phase approach to analysis is often termed 'higher order thematic analysis' and although the open coding level of analysis was adequate to serve the primary function of the data (i.e., item generation), these coded categories were subsequently grouped into broader domains to assist with the development of a conceptual framework of exercise-related life-quality. Higher order grouping is typical of thematic approaches to qualitative analysis and usually follows open coding (Holloway, 1997) and developing scale items from a conceptual framework is commonplace (e.g., Flanagan, 1978, 1982). Indeed Crone et al. (2005) have adopted a conceptual framework approach to help inform a psycho-social explanation of the physical activity and mental health relationship. For the purposes of scale construction, the development of a conceptual framework supports the process of the iterative assessment of content validity. A conceptual framework provides a reference point upon which to map items constructed for a new scale. Ensuring that each component of the framework is represented sufficiently by items increases the likelihood that the new scale will measure the underlying concept adequately.

The higher order grouping data fed directly into the subsequent scale construction phases of research therefore, some of the data included in the open coding level of analysis was not grouped into higher order themes as they would become redundant in the next phase of scale construction (Chapter III) and confuse the item generation process. For example, the term self-efficacy was not included in the groupings at this stage as this was considered too technical a term to form a question for the test measure. Negative language that, although was used by participants and was therefore suitably used verbatim for open coding but would not be suitable for question wording such as 'fear' was amended to 'worry'.

During this phase of research such amendments to language underwent an iterative phase of peer de-briefing (Spall & Stephen, 1998) so that a consensus between the research team was reached in terms of how best to represent what participants had reported in a manner suitable for item construction. Additionally, only those items that can be identified as causal or effect indicators are included in measures of life-quality (Fayers, Hand, Bjordal, & Groenvold, 1997) and so items of importance that did not reflect either a causal or an effect indicator of life-quality were also removed from the higher order coding and retained separately. In this respect the higher order category analysis also operated as a data cleaning phase before item design and generation commenced (Chapter III). To clarify the identification of causal and effect items, Streiner and Norman (2008) have explained that anxiety, for example may result in sweating, a feeling of impending doom, excessive worrying, irritability, sleep disturbance, difficulty concentrating and other symptoms. Streiner and Norman (2008) consider that such symptoms of anxiety should be considered 'effect indicators.' Similarly, Streiner and Norman (2008) have used QoL measurement to exemplify 'causal indicators' such as mobility, the ability to dress and feed oneself and participate in leisure activities. Figure 2 presents these examples provided by Streiner and Norman (2008) in the convention of structural equation modelling, whereby the construct is represented by a circle and the observed (or measured) variables are shown as rectangles.

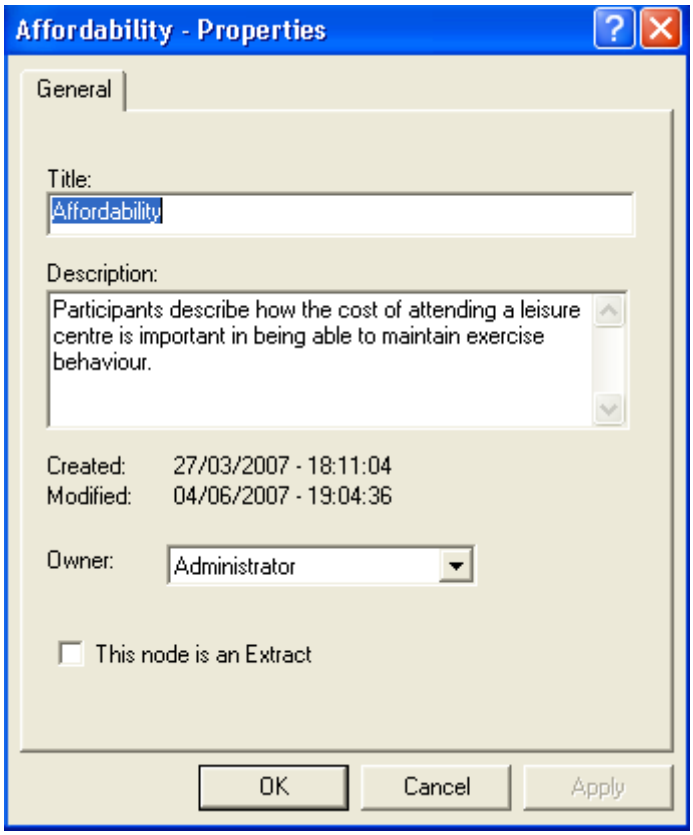
Figure 2. Causal and effect indicators (Streiner & Norman 2008).



QSR NVivo software was used to assist with the data management and analysis procedures and is a software package that has been utilised by other qualitative researchers (e.g., Bringer, Johnston, & Brackenridge, 2004, 2006). Transparency during the analytical and coding process was enhanced through the use of an electronic audit trail provided by the data development tracking functions of NVivo, including the use of screen prints and links with an electronic reflective diary (Bringer et al., 2004; Koch, 1994). In particular, the use of NVivo screen prints to be included as figures in a doctoral thesis has been encouraged (Bringer et al., 2004) to maximise transparency. For this reason, carefully selected screen prints are illustrated below to exemplify how NVivo was utilised as a research tool.

Nvivo allowed for the recording of written notes that helped to clarify the meaning of some of the codes (termed nodes by NVivo). For example, the node affordability was given some context to support the coded data (Figure 3).

Figure 3. Screen print of the description of the affordability NVivo node



Similarly, NVivo memos were made throughout the data coding process that allowed for elaboration upon the interpretation of data with respect to the research aims (Figure 4) and also (although beyond the scope of the current study) potential theoretical implications (Figure 5).

Figure 4. NVivo screen print of the memo indicating the psychological significance of better health.

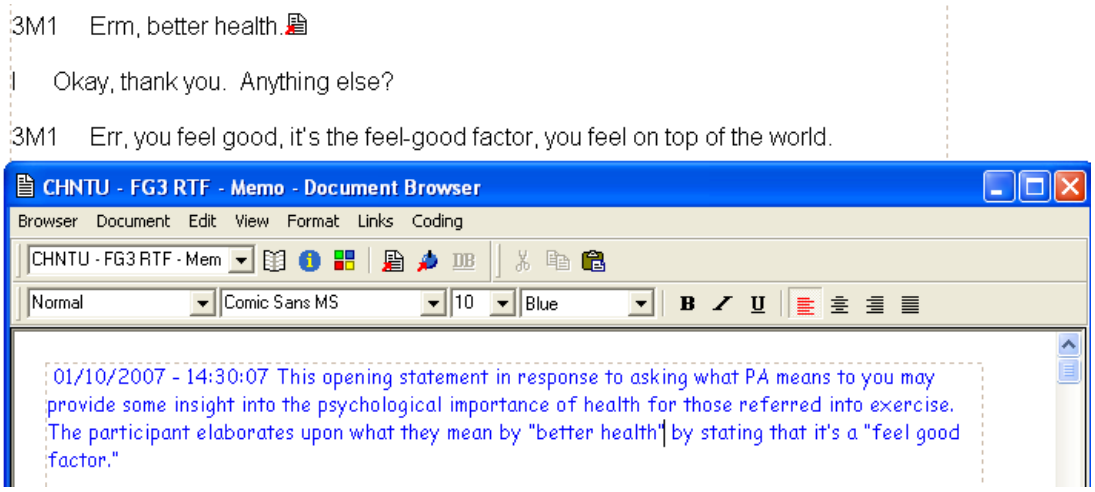
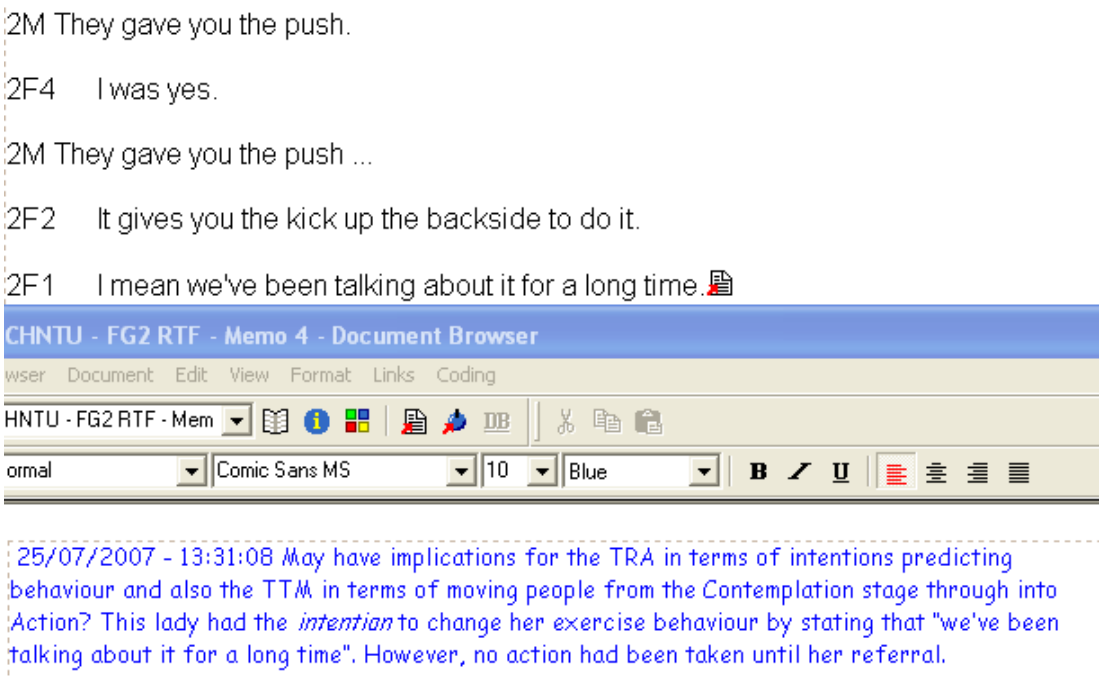


Figure 5. NVivo screen print of the memo indicating potential links between coded data and theoretical frameworks



The electronic research diary created in NVivo also allowed for the recording of potential research questions that while were beyond the scope of the current research would prove useful for future considerations (Figure 6). Additionally, the diary allowed for the recording of changes that were made to the coding of the data as the research progressed (Figure 7).

Figure 6. Screen print example of an NVivo electronic diary entry: Potential research questions

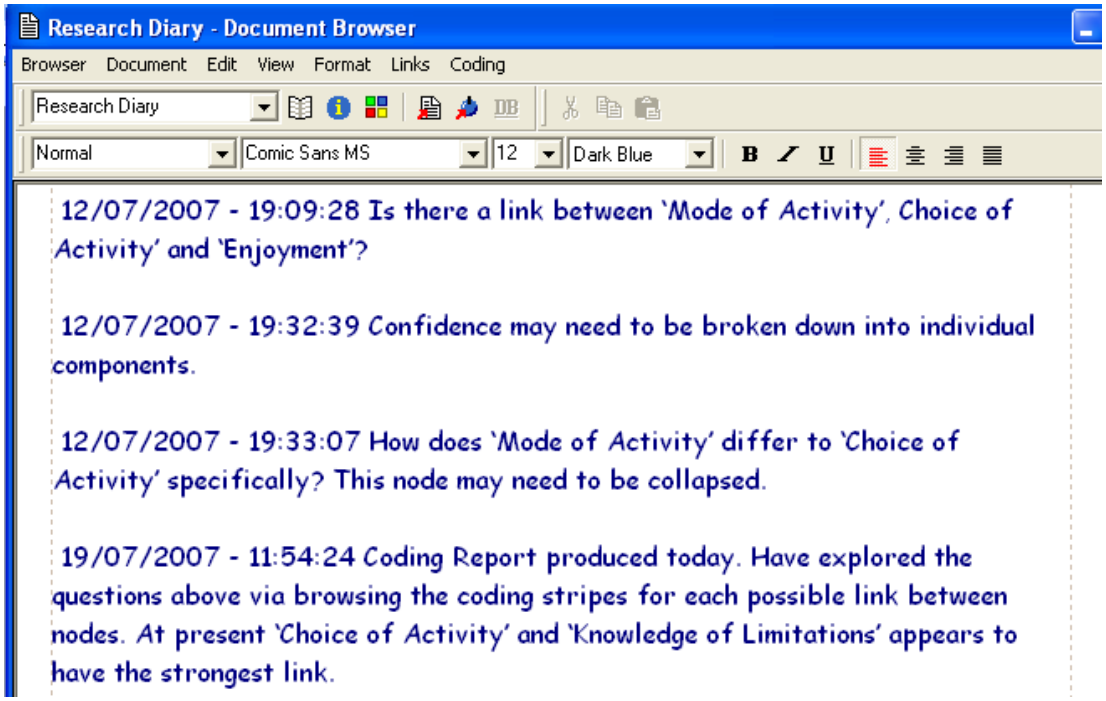
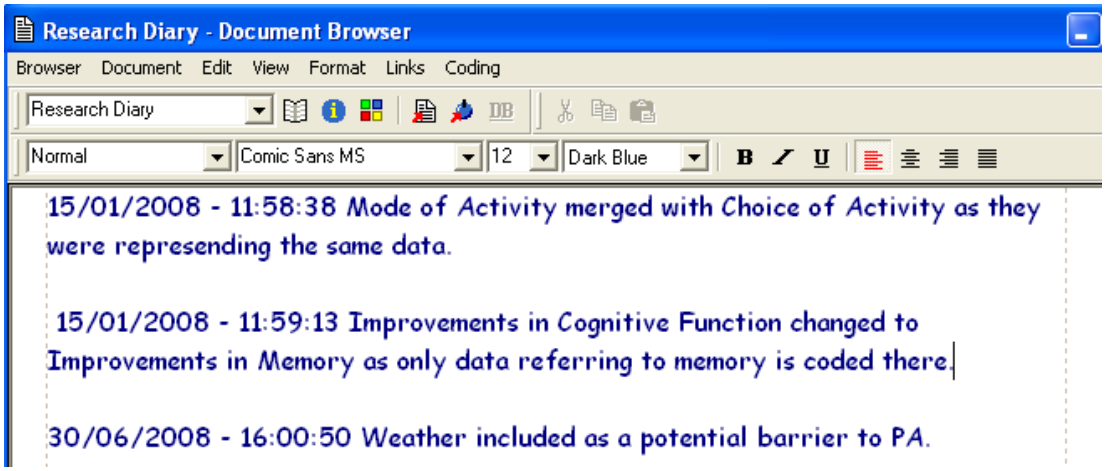


Figure 7. Screen print example of an NVivo electronic diary entry: Development of the coded data



Peer de-briefing (Spall & Stephen, 1998) was conducted with two independent experienced qualitative researchers who analysed approximately 25% of the raw data from each focus group. A meeting was held to discuss any differences in analyses and to reach a consensus whereby all analysts agreed on coding that represented the data in the most appropriate way. Such procedures have been noted as contributing to the credibility of qualitative research (Lincoln & Guba, 1985).

2.4 Results: Open Coding

While it is recognised that an open coding level of analysis is not often reported within the scientific literature - perhaps primarily due to word limitations, the open coded data is presented here for a number of reasons. Specifically, to maximise transparency (Bringer et al., 2004) and to demonstrate how the data generated from this initial phase of analysis a) translated into higher order themes, b) informed a contextual framework of exercise-related life-quality and c) shaped the next phases of scale construction. This general inductive approach to qualitative data analysis has been described as an easily used and systematic set of procedures that can produce reliable and valid findings (Thomas, 2006).

With the use of NVivo, the inductive approach to open coding level of analysis gave rise to a total of 70 nodes (or coded categories). Of these, 50 were free nodes and 20 were tree nodes (Appendix 4). Free nodes are categories coded by the analyst that do not contain sub-categories whereas tree nodes, as the name suggests are categories coded that have sub-categories attached to them. For example, the tree node 'Barriers to PA' had sub-categories that represented each of the barriers reported by participants which included: transport, time, childcare, poor health and the weather. Coding analysis of the five focus groups revealed codes that were considered important: a) before the start of an exercise programme, b) during the exercise programme and c) upon completion of the programme (12 weeks). Some of the coded data were also identified as being important in more than one phase of the exercise programme (e.g., before the start of an exercise programme *and* upon completion). Table 3 illustrates all the openly coded categories, when they were reported as being particularly important to the referral experience and indicates those categories that are shared with more than one stage. All of the coding categories revealed as important at each stage of the exercise programme including those that are shared will be reported, but only those categories that require clarification or that are of particular significance in terms of furthering our understanding of the relationship between exercise referral and life-quality will be elaborated upon with verbatim quotes. In instances where main coding categories have multiple sub-coding categories (referred to as tree nodes in NVivo), the sub-coding categories in the results that follow are indicated in brackets.

Table 3. Open coded data and the phase of exercise referral at which it was reported

Programme Start	Mid Programme	Programme End
*Affordability***	*Barriers to PA*	*Opportunity for PA*
Assistance with Surgery	*Using Gym Equipment*	*Changes in Energy Level*
*Choice of Activity**	*Helpful Leisure Centre Staff*	*Positive (Mental) Outlook*
*Choice of Time to Exercise**	*Achievement****	*Self-Esteem*
Exercise Referral to Initiate PA	*Comparisons to Others*	*Enjoyment of PA*
Expectations of the Gym Environment	*Competitiveness/ Determination****	*Feeling Healthy (Mentally & Physically)*
Fear of Exercise	*Good Instructors*	*Feeling Younger*
Fear of the Gym		*Social Support*
Fear of Injury		*Making Time for Self*
History of PA		*Feeling Comfortable with Self*
*Good Instructors (knowledge/support/ communication/familiarity)*****		*Feeling Better/ Feeling Better about Self*
		Improved Physically Active Lifestyle
		Improved Sleep
		Improvements in Confidence and Self-Efficacy
		Enhanced Self Awareness
		Improvements in Independence
		Improvements in perceived QoL
		Improvements in Memory
		Improvements in Motivation
		Improvements in Social Contact
		Improvements in Stress Management
		Managing Depression
		Family Time
		Absence of Pain
		Pain Management, Weight Management
		Healthy Eating Habits
		Disease Management
		Perceptions of Life Expectancy
		Knowledge of Illness
		Knowledge of Limitations
		Knowledge of PA
		Maintaining Exercise Behaviour
		improvements in ADL

Note. *Represents coded data reported at the start and during exercise referral. **represents coded data reported at the start and upon completing exercise referral. ***represents coded data reported during and upon completion of exercise referral.****represents coded data reported at the start, during and upon completion of exercise referral. Active daily living is abbreviated to ADL.

Codes that were considered important before the start of the exercise programme

Prior to the commencement of an exercise programme, participants reported the following areas of importance/relevance: Assistance with surgery, exercise referral to initiate PA, comparisons to others, expectations of the gym environment, fear of exercise, fear of the gym, fear of injury and history of physical activity. Affordability was reported as important both at the start and upon completion of the programme and is therefore reported in a later section below.

Assistance with Surgery

Participants reflected on their understandings of how PA and exercise may benefit various surgical procedures. For example:

"So when you have the operation, whatever it is whether it's a trunk operation or a leg operation or arms or whatever it is, it builds up that muscle and takes away the fat, which is easier for the surgeons to operate you on."

"Erm....It helps your heart so when you're put under anaesthetic or anything like that you can withstand the anaesthetic. Erm it's ... and when you come out of hospital, one you've already taken that step and you're fitter, you get better quicker."

"Erm now I feel less worried about going to have an operation now I've got fitter.... I was worried about my lung capacity, the heart capacity and the fitness of my complete body before I went for an operation. Now I'm not worried."

Exercise Referral to Initiate Physical Activity

For a lot of respondents across the five focus groups, a referral for exercise gave them the opportunity to put into action, growing intentions to exercise that had not been acted upon previously. For some, exercise referral provided the opportunity to initiate a physically active lifestyle for those that had given it little consideration. For example:

"and I thought this is going to be the kick-start I need."

"and if you've been ill and you have been sedentary for a long while ...

Interviewer Mmm.

... it's that first move to actually get you off to your feet ..."

"I think it's just the opportunity, it's the opportunity to actually go and do it. You know, coming from a background of not having any sort of exercise like that before, it's just given me that push to go and do it because I wouldn't have thought about going to do anything like that."

Comparisons to Others.

Participants reported their reflections of comparing themselves to others before the programme started. Some examples included:

"I was beginning to think I'm fat, I'm a blob, I'm unfit and I'm going to be going to a gym that's full of super fit people ..."

"They don't think about your cellulite or what you look like. They might think 'crikey, I'm slim in comparison to them that's just walked in' or 'is she really fatter than me or am I fatter than her?' You know, it's all that self-comparison."

"It's a fear of what do I look like, you know."

"And me in my 40s thinking 'I shouldn't be here.' And I was ... I walked in there really, really scared."

However, these self-comparison concerns were reduced once the participants had the opportunity to experience the leisure environment. For example, participants commented:

"You think people are going to laugh at you ... well that's what I felt. You know, and as he says, nobody's really interested in what you're doing ..."

"Well I did find... when I first started and this is when I broke my leg, the scarring worried me, my size worried me.... and in the end I started swimming after the ...

going in the gym and I couldn't care two hoots then. Once I'd done ... took that first plunge so to say in the water."

"But that's how I find it, that it's the self-worry about what other people think about. I mean nobody cares less really."

One male participant asserted that the pre-exercise self-comparison concerns reduced completely once his programme started:

"Well once you get into it, that [worry about comparison to others] disappears doesn't it?

Expectations of the Gym Environment.

Prior to the start of the exercise programme, participants expressed initial concern regarding their expectations of the gym environment. For instance:

"Erm, very apprehensive, I didn't know what was going to happen. I'd never been in a gym in my life."

"I went, I got this appointment with this [name of the instructor]. I mean I had no idea who [name of the instructor] was and I went absolutely terrified."

However, this initial expectation changed once participants had the opportunity to experience the leisure environment. For example, one female participant recalled how her expectations of the gym environment changed after her first visit:

"But then when I walked in, there's all different people, all different sizes and they were all friendly, didn't matter whether they were larger or smaller. But for me, I just felt as though because I was ... there was a lot of mature ladies there, a lot of the older ones, which really opened my eyes to see probably some of them in their 60s and 70s in a gym."

Fear of exercise.

In some instances, a genuine 'fear' of exercise was reported. For example, participants indicated:

"For me, I didn't do any exercise whatsoever and exercise was a fear, a huge fear."

"I think it was a fear of it really because I've got a back problem, you don't know what you're doing, you don't know how much to do."

However, again this fear lessened once participants had the opportunity to experience exercise at the leisure centre. Participants reported:

"I got in and I was frightened to death, I had to have a waddle to hold me up and I just had to come back and just swim to the edge. But by the time I'd finished that first session … I know it sounds daft; I'd done six widths."

"And once you know that your body can do it, you're not frightened to carry on and get better with it."

Fear of the gym.

Participants who were referred to a gym were more specific regarding their pre-exercise programme concerns. For example:

"I was scared."

"I was frightened to death. (laughs) I dreaded going."

Again, these feelings changed once the participants had experienced the gym environment. For example:

"Well I was getting butterflies before I went and my husband would say 'don't start, you've got butterflies, you want to go to the toilet', I said 'oh I know but once I get there I'm alright'. And I was, as soon as I walked in there and saw [one of the instructors], I was fine; that's stopped now, that doesn't happen to me at all now, I look forward to going. Whereas I was in a way I suppose in my mind, dreading walking in."

Fear of injury.

Participants also reported pre-exercise programme concerns regarding injury. For example:

"So I thought to myself I'll be a lot better in a gym, where I can go on a treadmill walking and not be afraid of falling because falling is a big dread that I have."

"Yeah because there is a danger of doing too much too soon."

History of Physical Activity.

Some of the fears and concerns reported may have been mediated by a participant's history of PA. For example, one male participant stated:

"Maybe my experience is different isn't it because I mean I belonged to a gym before, so I was familiar with the machines."

Other reports included reflections of a more active history that participants were hopeful to regain. For instance, one female participant recalled:

"I used to go a long time ago to a gym and I've always missed it and it was a ... a reason to start going again."

Codes that were considered important during the exercise programme

Once individuals were engaged in an exercise programme, there were factors that were considered important for the enjoyment and maintenance of exercise behaviour. Some areas overlap those described above (before the start of the exercise programme), and some are unique to mid-programme participants. Choice of activity and choice of time to exercise were reported as important both at the start of a referral programme and during and are reported in a later section below. Those areas that were reported as unique during the exercise programme are reported here and included: recognition of the potential barriers to PA (e.g., transport, time, childcare, poor health and the weather), using gym equipment, helpful leisure centre staff, achievement and good instructors. Data coded as competitiveness/determination was an area of importance reported as relevant both during and upon programme completion and is therefore reported in the relevant section below.

Potential Barriers to Physical Activity.

As participants reflected upon their progression through the 12-week scheme, potential barriers to PA were identified. These included: (a) transport, (b) time, (c) childcare, (d) poor health and (e) the weather. Some examples are as follows:

"It's hard for them [people participating in the exercise programme] to get here as well." (transport)

"And it's just … it's also down to time management isn't it, fitting everything else in a person's life. I work, I have a family you know, I have family living out of Nottingham and it's trying to take on board when can I fit these things in." (time)

"And I know for me I had to miss two weeks of my 12 weeks because I had childcare to think of over Christmas. So that was a shame, I didn't want to miss it but I had to, that's another thing." (childcare)

"I've been sick for the last couple of weeks, I want to go back up there Monday, nine or ten o'clock and I'm going to see [instructors name] and get cranked up again and go for some more, you know." (poor health)

"And particularly in the cold weather because this complaint that I have it's called … well it starts here in the temple but it causes glaucoma." (the weather)

Using Gym Equipment.

Reflective accounts of using gym equipment fell into two main categories: a) the benefits of using gym equipment, and b) the challenges of using gym equipment. For example, the benefits of gym equipment were described in the following manner:

"So that's why you need these special machines because these machines you can sit on them and get off them."

Challenges of using gym equipment were reported more frequently, for example:

"Because sometimes, pressing this button and pressing that button, and I used to get all mixed up and all terribly confused."

Helpful Leisure Centre Staff.

Participants described the importance of having staff that were described as helpful throughout the leisure centre they attended and not just in terms of their experiences of the instructor with whom they worked. For example:

"The reception staff were very good to start off with, they said where to go and gave me a map and stuff and I went up."

"I think the staff's a big help as well because they're always so welcoming."

Codes that were considered important upon completion of the exercise programme

There were a wealth of codes that were considered important upon completion of the programme. Affordability, achievement and competitiveness/determination were reported as important both during the exercise referral programme and upon completion and are therefore not elaborated upon here. Again, in the main these codes differed in terms of the context in which they were reported. For example, affordability was discussed as an important consideration for participants with respect to initiating attendance at paid activities *before* the start of a programme of activity and also in terms of being able to maintain their attendance *upon completion* of an exercise programme.

There were also a number of areas unique to this phase of referral including:

- opportunity for PA
- changes in energy levels
- positive (mental) outlook
- self-esteem
- enjoyment of PA
- feeling healthy (mentally & physically)
- feeling younger
- social support
- making time for self
- feeling comfortable with self
- feeling better/feeling better about self
- improved physically active lifestyle

- improved sleep
- improvements in confidence and self-efficacy
- improvements in independence
- improvements in perceived quality of life
- improvements in memory
- improvements in motivation
- improvements in social contact
- improvements in stress management
- managing depression
- family time
- absence of pain
- pain management
- weight management
- healthy eating habits
- disease management
- perceptions of life expectancy
- knowledge of illness
- knowledge of limitations
- knowledge of PA
- maintaining exercise behaviour
- improvements in active daily living (ADL).

All the areas uniquely reported as important at this stage of the referral programme are supported by verbatim quotes save for improvements in sleep and memory as it was considered that these required no further clarification. Additionally, the notion of being 'comfortable with self' was commonly reported as being linked to perceptions of self-esteem and so for clarity this is reported under 'improvements in self-esteem.'

Opportunity for Physical Activity.

The perceived decrease in the opportunity for PA was largely discussed as a more general difficulty and some participant reports attributed the decrease in active daily living to advances in technology. For example, participants reported how the use of computer aided technology such as e-mail omits the necessity to walk to the post box to post a letter:

"There's no writing letters. And even posting it, you don't have to post it."

Changes in Energy Levels.

Participating in a programme of exercise impacted upon perceived energy levels in different ways. Some participants reported an increase in perceived energy levels, while others reported a perceived reduction. The majority of reports described perceived increases in physical activity (22 reports) as opposed to perceived decreases (7 reports).

Positive (Mental) Outlook.

Several references across the focus groups were made with respect to exercise facilitating a positive mental outlook. This was often explained in terms of general outlook on life and the perception of the ability to cope with stressful situations. For example:

"Well just more positive about most things."

"It affects your mental ability to cope with life."

Improvements in Self-Esteem.

Many participants across the five focus groups reported improvements in self-esteem. This was attributed to perceptions of improved psychological health. In particular, improvements in confidence related to perceptions of self-esteem. For example:

"Well, just how I feel about improved health, which then has a knock-on effect to you know, so many parts of my life. As we said, confidence, self-esteem."

The notion of 'feeling better about self' was also described in terms of improvements in self-esteem. When describing perceptions of improved self-esteem, one participant reported:

"Yeah. Just feel better in … feeling better in yourself, you know."

This is particularly salient given that exercise referral practitioners frequently attempt to deconstruct the perceptions that contribute to the often reported 'I just feel so

much better' statements made by those participating in schemes. This data suggests that reports of feeling better could be directly associated with improvements in perceived self-esteem.

Enjoyment of Physical Activity.

Reports of enjoying physical activity were made, and for some, this new-found enjoyment was unexpected:

"I would never have believed it but exercise for me has brought pleasure really, yeah. I really do enjoy it."

Feeling Healthy (Mentally & Physically).

General good health was often reported as a combination of both positive mental and physical perceptions.

"Even my blood pressure's down, which is smashing." (physical)

"It gives you that well being feeling when you've exercised." (mental)

Feeling Younger

Participants reported that their referral for exercise had made them feel younger. For example:

"It [exercise referral] *makes you feel young again as well."*

"I mean you don't feel 21 again but you do feel young again."

Importance of Social Support.

Sources of social support, particularly from fellow participants were reported:

"I suppose that's the case of everybody going at the same time and then meeting each other. If I go on my own now, which I'm going on my own now, it doesn't feel the same."

Improvements in Making Time for Self.

Newly developed exercise behaviours were reported as providing opportunities for time away from others. Participants reported valuing the time to spend on themselves:

"I find it's time for me as well, it's that me time that I'm doing it for me."

"I think one thing I find, I tend to do things for me now, not for anybody else, I just do things for me. You know, I've always done things for other people and now I think well I come first now, my turn."

"Yeah, it's 'me time'."

Feeling Better and Feeling Better about Self

A number of reports were made that used language stating that a referral for exercise had made participants 'feel better' or 'feel better about themselves'. While probes had helped to clarify more specifically what was meant by these statements and in what way, some of these reports were difficult to deconstruct so were simply coded using the participants language (i.e.,, feeling better and/or feeling better about self). For example:

"Feeling good when you've done it. I think you feel a lot better when you've done exercise."

"I may not look any better but I feel better when I come out. You're taller, you've got a spring in your step."

Improved Physically Active Lifestyle

Many participants described how exercise referral had impacted upon other areas of their life to help them incorporate PA into their daily routine. For example:

"I've found that since I've been doing exercise I either get my bike or I walk. If it's within that kind of range you know, I don't go crazy like you know, instead of where I normally may have taken the car a mile, I'll walk it now."

Improvements in Confidence and Self-Efficacy

Participants were able to identify types of confidence, namely: a) self-confidence, b) confidence in physical ability, c) confidence with gym equipment and d) confidence in active daily living (ADL):

"So I feel more confident in myself, a lot more." (self-confidence)

"And by you know, doing certain exercises, taking it easy at the start and that gave me the confidence to sort of know that I could move around." (confidence in physical ability)

"And confidence in the gym, I never thought I'd be confident in a gym and I now know I could walk into a gym anywhere in the country and I ... or anywhere in the world, and I'd feel happy because I know the machines, I know what does what and I could just get on with it which is big thing really." (confidence with gym equipment)

"It's just given me more confidence. If anybody used to ask me if I wanted to do something at work, I used to sort of shy away from it, 'oh I can't do that, I'm not capable of doing that'. But now, yeah I'll go for it. Okay, so I might finish up with egg on my face but nevertheless I will have the guts to get up and have a go at it. And virtually everything I've tried so far I've got to the end and I've succeeded. Sometimes it's been slower than others but it gives you the confidence to go out and do it." (confidence in ADL).

"I've been surprised with myself that you know, during those 12 weeks what I was able to do on all those machineries." (self-efficacy).

Improvements in Independence.

Exercise was described as supporting participants to maintain an independent lifestyle and in some instances, the role of exercise was reported as helping to regain lost independence following ill health:

"Your independence goes but once you've got on your feet and you get going, you don't get frustrated anymore, you're part of the community, you're part of your family, you get your independence back, you get your happiness back."

Improvements in Perceived Quality of Life.

Participants were able to report the impact of their exercise programme upon their perceived quality of life (QoL). Descriptions of how exercise had altered personal perceptions of the concept of quality of life were offered and were sometimes unexpected as a health outcome. Indeed, one participant described how they separated the notion of health improvement from QoL, although they recognised the holistic nature of their responses to exercise referral:

"I just thought the health would improve but a lot more improved….quality of life and everything. You know, it all joined up."

Improvements in Motivation.

The source of motivation to exercise was often described in terms of the activity itself perpetuating the desire to maintain the active behaviour. A referral into exercise helped to break the cycle of sedentary behaviour and offer opportunities for a more physically active lifestyle. Physiological rewards were attributed to initialising motivation. For example:

"So I thought I'd have a go. And for me, it did what I wanted it to do; I lost the weight. It gave me motivation, I got ready to go, I got … you know, I got the incentive to go, which I think you need."

Additionally, psycho-physiological mechanisms for the source of motivation experienced by participants were offered. For example:

"It's probably a cycle that you get into, that the exercise starts working on the brain chemicals, the endorphins and you feel better about yourself, you get more motivated."

Improvements in Social Contact.

A referral into exercise proved to offer opportunities for social contact with others, and for some, new friendships were formed (e.g., "and I've made a lot of friends").

And for others, social contact developed into general conversation with others referred:

"You find after a few weeks, the same people are going at the same time and you start nodding, then you say hello in the changing room and just general chat starts up."

Improvement in Stress Management.

Participants acknowledged the role of their exercise programme in helping them to manage potential stress.

Improvements in the Management of Depressive Symptoms.

Participants recognised the mood-enhancing effect of exercise, particularly in the management of depressive symptoms. For example:

"Yeah. It's lightened my mood, I'm not so … I did suffer from depression actually, so … I'd forgotten about that actually. So yeah, that's really … that's improved."

Family Time.

Time spent with family members was considered important for those completing a 12-week exercise programme. The goal of maintaining adequate family time was often an important motivator to maintain exercise behaviour:

"Being able to do the things that you used to do, well again. Not being frustrated because you can't do them and you feel like part of the family again because you can actually do things with the family like you used to do before."

Absence of Pain.

For those participants who had been referred into their exercise programme to assist with pain management, some reported a complete absence of pain as a result of completing their 12-week programme (e.g., "I'm just pain-free"). These reports were distinct from those that made reference to a reduction in the experience of pain or changes in perceptions of pain management. Therefore, data were coded into two distinct areas of pain experience: absence of pain and pain management. Doing so also ensured that the experience of pain was represented adequately and accurately during item generation.

Pain Management.

In the main, for those who did not report a complete absence of pain as a result of their exercise programme, exercise was still recognised as an important method of managing symptoms of pain experienced by participants:

"I get niggles, don't get me wrong, it's not disappeared but it's just manageable now. Compared to what I've had, it's fantastic."

Weight Management.

Weight management was described in terms of weight loss, weight gain or weight maintenance:

"And then it'll start coming off again. And it does and then I go on holiday and eat and drink too much and put it back on again." (weight loss & gain)

"I needed to put weight on." (weight gain)

"Yeah, yeah, I mean you know like you go out and buy clothes and you think oh nothing fits me, I'm not going up another size, I'm going to manage with what I've got." (weight maintenance)

Healthy Eating Habits

Participants reflected on the importance of healthy eating to facilitate their exercise programme and any other additional physical activities undertaken as part of their new healthy lifestyle. For example:

"Because it's exercising and diet at the same time."

"I mean both with eating and with physical exercise."

"It's [exercise referral] *made me think more about healthy eating, without a doubt."*

Disease Management.

The role of exercise was recognised as important for future health and as an important mediator for disease management and the likelihood of poor health. For example:

"So really [the name of the exercise referral scheme] *in that way, it's not just a positive move to make you do more later ... but a positive move to make you healthier and to stop you from ever going back into the hospital for whatever reason it might be."*

Perceptions of Life Expectancy.

One participant who had been diagnosed with a terminal illness reported that the main outcome of his referral into exercise had improved his optimism for life expectancy and developed the ability to respond to his new found optimism:

"Being positive about life expectancy, I think that's probably the main one. And knowing that having got that life, I have the stamina to see it through."

Knowledge of Illness.

Participants made several reports of their knowledge of the illness with which they had been diagnosed. For example:

"Now if you've got an under-active thyroid gland and a broken leg, you also get sometimes arthritis in the leg that you've broken."

Knowledge of Limitations.

Knowledge regarding a participant's illness and limitations played a key role in their perceptions and confidence of what physical activities would be attempted:

"But if you're certain weight or you've got a certain disability, once you get down on that floor, you can't get up."

"And I have to take it slowly."

Knowledge of Physical Activity.
Participants made frequent reports regarding their understanding of the benefits of PA:

"Apart from your heart rate and your general well being, it increases your stamina a bit. Well I found it did anyway."

These reports included references to the benefits of the different types or 'modes' of activity available as part of the exercise programme. For example:

"It's in water, so the water's taking the weight off your body."

Improvements in Active Daily Living (ADL).
Participants were able to recognise how an increase in PA had impacted upon other areas of their general lifestyle and improved their ability to undertake everyday tasks. Examples included:

"It's helped me a lot, really a lot, not just a bit. I mean my husband will tell you, it'd take me two hours to clean three rooms, now it's in less than an hour and I'm doing them well."

"Yeah, gardening; I do all my own seeds and my baskets and pots. I've made over 30 pots and you know, two years ago … well a year ago rather, I couldn't have done that."

"Oh no but I've never been in bed past six and now I'm back to that, then I get up, do breakfast you know, for the wife and what have you, take the dog for a walk. Stupid things you know, only little things but it adds up."

Maintaining Exercise Behaviour.
A number of affirmations and expressions of intentions to maintain newly developed exercise behaviours were reported:

"But I wouldn't give it up for nothing."

"That's it you know because you don't think about the exercise, you go and do it."

"You know, it's not … it's not finished once the 12 weeks are over; I'm carrying on going."

Codes that were considered important at the start of the programme and upon completion

Affordability.

The ability to afford leisure services was reported as something to consider early on in a participant's referral. This consideration increased in importance as the referral period progressed because, as with many schemes, the cost of leisure services increases once the 12-week intervention phase has been completed. Although participants are offered a long-term subsidy, many individuals were mindful of considering the overall expense of participating in structured exercise. As a consequence, affordability was reported as important both at the start of an exercise referral programme, and upon programme completion. For example, participants explained:

"Well if somebody has done three months on their £1.50 and they're doing okay and suddenly it goes up to … what is it, £3.20 or something, and that person is on a limited budget, then the chances are they're going to have to stop going."

Codes that were considered important at the start and during the exercise programme

Choice of Activity.

In the main, codes that were reported as important to participants both at the start and during the exercise programme differed in terms of the context in which they were reported. For example, choice of activity played a key role in the likelihood of a participant attending initially (important before the start of an exercise programme) but was also an important factor in mediating the likelihood of subsequent attendance (important during the exercise programme).

The choices regarding the type of activity that were undertaken appeared to play a role in four key areas, namely: (a) the likelihood of starting an exercise programme,

(b) enjoyment, (c) supporting symptom management and (d) the likelihood of maintaining exercise behaviour.

Participants felt that it was important to have a choice about which activity they undertook; e.g., some participants indicated the following:

"I must say … I have to say I did have a choice, the nurse at the practice who referred me, she said 'what would you like to do and what can you do?" (likelihood of starting programme)

"But I find the treadmill, I love it, I could stop on that all the time, I love that." (enjoyment)

"I mean I walk quite a lot….but I enjoy going to the gym, it's something new for me, I've never been to a gym in my life." (enjoyment)

Different female participants explained why having a choice of activities to choose from during their referral period was particularly beneficial:

"But the thing is with bad knees, you don't put that pressure on your knees, you don't put that pressure on your hips or your back." (supporting symptom management)

"Because I used to do aerobics; well I can't do that now. And I couldn't do Pilates because I can't get up and down from the floor, the same with yoga. So she said 'well would you like to try the gym?' It was a choice." (supporting symptom management)

Participants often reported trying different modes of activity as a result of a positive experience undertaking the activity for which they were referred. In turn it appeared that this willingness to undertake other modes of PA had a direct impact upon the likelihood of them maintaining their newly-adopted exercise behaviour. For example, a female participant stated:

"So I am just going to try and work at it and chuff away at it but that's the only thing I can do and maybe try different things, I keep trying different things. You know,

exercise or walking or whatever and go on from there." (the likelihood of maintaining exercise behaviour)

Choice of Time to Exercise.

In addition to choices regarding the type of exercise undertaken, choices regarding the *time* at which participants engaged in their exercise was also important for enabling people to start their programme of exercise. Quite simply, if a participant was unable to attend an exercise appointment they were unable to start their programme. For example:

"Well I'm at work anyway but I could have chosen any time…But I mean I can go more if I want, although I'm at work during the day and that's why I go Saturdays and Sundays." (choice of time to exercise).

Codes that were considered important both during the programme of exercise and upon completion

Achievement.

Participants who reported feeling that they had achieved something during their referral into exercise described the importance of successfully reaching targets or goals. This was important not only to help participants finish their 12-week prescription but to continue exercising thereafter. For example:

"I set myself targets. And I reach them and when you reach your target you feel, you know, really great."

"The one thing I ever wanted to do was dance at my own son's wedding after I broke my leg and I managed to do it because of [the name of the exercise referral programme].

"And it's me setting goals for myself, now and in the longer term it's all part of it."

Competitiveness/Determination.

Individuals who identified that they had a competitive nature reported that a referral for exercise had re-instated their competitive behaviours:

"It's made me competitive again. Before, when I wasn't fully fit, you couldn't compete at the same level but now I can."

Equally, participants recognised that their determination had played a key role in their ability to complete the initial 12-week referral. For example:

"So I was quite determined that there was nothing going to stop me."

"Because I'm determined not to give in and carry on."

Codes that were considered important at the start, during and upon completion of the exercise programme

Good Instructors.

Often, one of the attributions participants made to their positive experience of exercise referral was their positive perception of the instructors with whom they worked. Although in most cases, these positive attributions could be further coded into the importance of instructor: a) knowledge, b) support,
c) communication and d) familiarity, some participants simply referred to good instructors without any elaboration (e.g., "they [the instructors] are very good").
For this reason, this data were coded simply as 'good instructors'.

There were several examples of specific factors related to good instructors that were also reported. In addition to identifying the specific attributes of the exercise instructors that were important to the referral experience, participants were also able to communicate that the role of the instructor was important at the start, during and upon completion of their programme. Instructor attributes were not matched to any stage as such although in some examples below verbatim quotes are used to illustrate both points:

"I was always frightened if I went to a gym without the…you know, sort of [the referral scheme]*, that the person teaching me wouldn't know what they were doing and I'd finish up in a worse state than ever."* (knowledge/ importance of the instructor at the start of the programme)

"[one of the instructors] again, as everyone's said, was brilliant. And erm, I think I had a couple of sessions with her and I always say hello to her now and just … she says, 'are you alright?', 'yeah, yeah', as I'm carrying on on the cross trainer, you know." (laughs) (importance of the instructor during the exercise programme)

"Because I need to know, with not being in a situation like that before, in a gym before, I still need to know what I'm doing after the 12 weeks. You know, so they don't leave you alone, it's … you get extra help if you want it." (importance of the instructor upon programme completion)

"Well I think … yes, I think after the 12 weeks they're still supportive, very supportive and they're there all the time to give you as much help as they can." (support/ importance of the instructor upon programme completion)

"They seem to be able to communicate with you on your level." (communication)

For some participants having a variety of instructors to work with was a positive experience, and for others it was important to work with the instructor who supported them with their initial appointment. For example, participants who enjoyed working with a variety of instructors reported:

"I've just had a bit of variety of them all the time and they were all very, very helpful." (familiarity).

While others reported the importance of being important to feel more familiar with the instructor and work with the same one wherever possible. For example:

"Oh well I didn't and I found that a bit disheartening because you got used to one and then I was …I got put with somebody else." (familiarity).

2.4.1 Results: Higher Order Domains

The second phase of analysis was to translate the open coding level of analysis into higher order categories or themes. This allowed for the refinement and linking together of the openly coded data and followed the process suggested by Rubin and

Rubin (2004) whereby the coded data were assessed to examine where commonalities existed and were grouped together. These smaller groups of themes were then assigned a title that reflected the coded data contained within it. Again, this process underwent a series of iterative peer de-briefing (Spall & Stephen, 1998) whereby members of the research team offered suggestions for amendments to the higher order theme titles so that the most fitting description was used and a consensus between the research team was reached.

The higher order categories that comprised the framework were: leisure facilities and lifestyle physical activity, confidence and motivation, discomfort and fitness and mental and physical well-being (Tables 4, 5, 6 and 7). This phase of analysis assisted with the development of a conceptual framework of exercise-related life-quality (Figure 8) which would inform the later item generation and analyses phases of scale construction. Those codes/themes that were not suitable for item construction due to inappropriate language (e.g., jargon or too vague) or that were not causal or effect indicators of life-quality (Fayers et al., 1997) can be reviewed in Table 8.

Table 4. Coded categories grouped into the domain of leisure facilities and lifestyle physical activity

Leisure Facilities and Lifestyle Physical Activity Domain
Cost of leisure facilities
Choice of type of exercise
Choice of time to exercise
Perceived barriers to exercise (transport, time, childcare, poor health and the weather)
Enjoyment of physical activity
Social contact with others
Weight management
Incorporating physical activity into daily lifestyle
Healthy eating habits
Competitiveness
Determination
Knowledge of the benefits of physical activity

Table 5. Coded categories grouped into the domain of confidence and motivation

Confidence and Motivation Domain
Confidence using gym equipment
Confidence exercising in a leisure centre with minimum support
Confidence to participate in regular physical activity and exercise
Confidence to undertake daily non-physically active tasks
Confidence around others
Confidence in physical ability
Confidence to join in activities with family and friends
Perceived importance to maintain a physically active lifestyle
Confidence to maintain a physically active lifestyle
Motivation to maintain a physically active lifestyle
Likelihood of maintaining a physically active lifestyle
Motivation to be physically active

Table 6. Coded categories grouped into the domain of discomfort and fitness

Discomfort and Fitness Domain
Frequency of injury preventing physical activity
Frequency of physical pain
Frequency of pain symptom management
Frequency of illness symptom management
Frequency of physical fitness achievement

Table 7. Coded categories grouped into the domain of mental and physical well-being

Mental and Physical Well-being Domain
Comparisons to others while exercising
Worry regarding exercising in an environment with others
Worry regarding participating in structured exercise
Worry regarding the prospect of exercising in a fitness suite/gym
Frequency of the opportunity to make time for self
Perceived energy level
Outlook on life
Perceived self-esteem
Perceived physical health
Perceived mental health
Sense of well-being
Quality of support from leisure staff
Quality of support from friends and family
Sleep quality
Memory function
Stress management
Mood management
Weight management
Perceptions of life expectancy
Active daily living
Overall quality of life rating

Figure 8. A conceptual framework of exercise-related life-quality

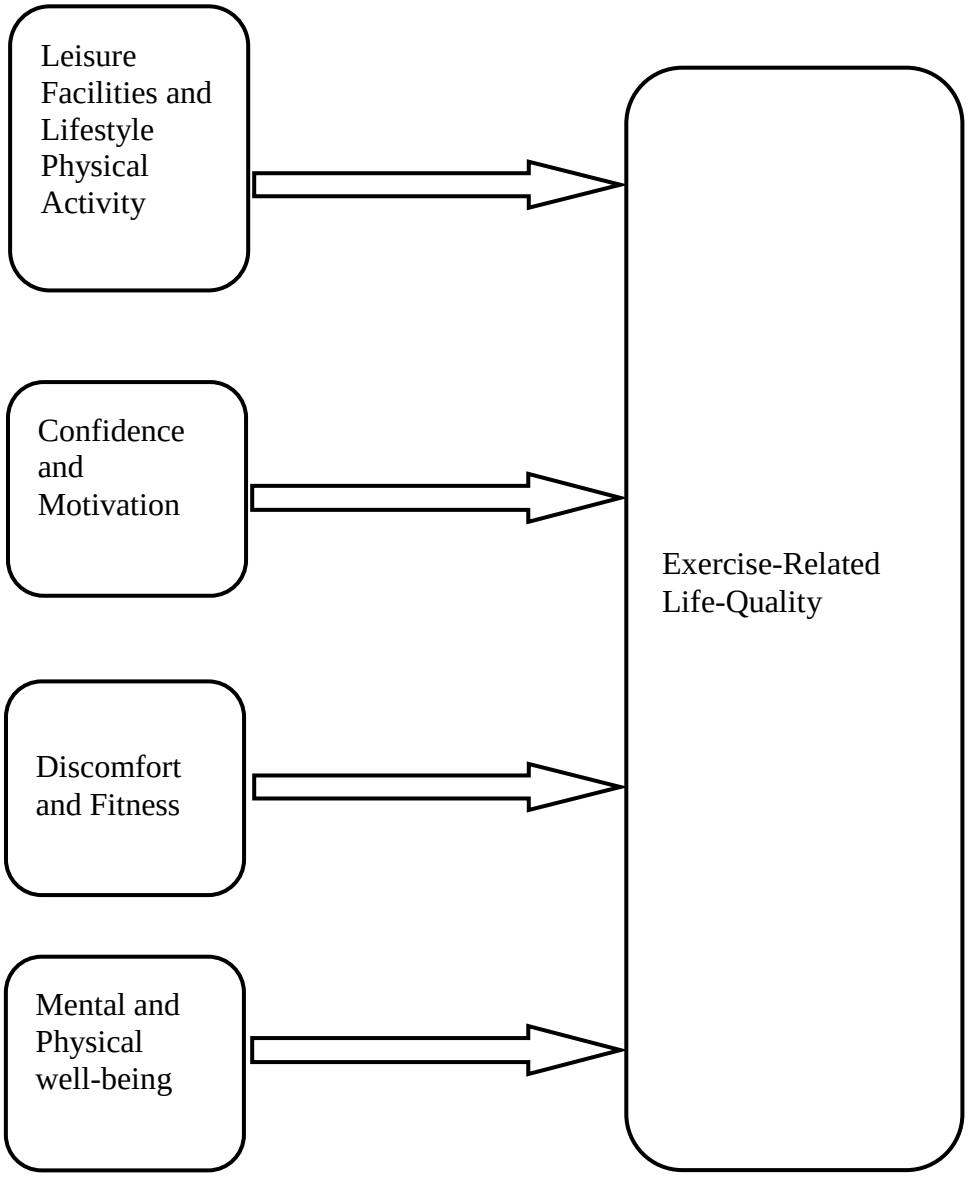


Table 8. Data not included in the higher order analysis and retained separately

Coded Category	Reason for Exclusion
Assistance with surgery	Not a QoL cause or effect item
Exercise referral to initiate PA	Not a QoL cause or effect item
Feeling Better	Too vague
Self Efficacy	Jargon
Self Awareness	Too vague
Good Instructors: Importance of knowledge	Not a QoL cause or effect item
Good Instructors: Importance of Support	Not a QoL cause or effect item
Good Instructors: Importance of Communication	Not a QoL cause or effect item
Good Instructors: Importance of Familiarity	Not a QoL cause or effect item

2.4.2 Reflections of the Interviewer

I found the questioning route particularly useful to help guide me through the areas of discussion but I also found the structure of introductory, transition, key and ending questions helpful in terms of setting the pace at which the topics were addressed. Across all five groups I did not sense that participants felt rushed in their efforts to respond to questions and probes, despite my being aware of keeping the interviews to a desirable time of around one hour. I also noted that participants were equally able to describe their experiences of exercise-related life quality in general terms in addition to providing more personal accounts. I think that taking the time to 'set the scene' by the informal and inclusive way in which the aims of the study were communicated to participants alongside my role in the research contributed greatly to the development of an early rapport with participants. The NHS template information sheet that I was required to provide participants with was quite lengthy and I felt that this could potentially challenge this rapport so I was grateful of the shorter summary sheets that I had prepared. Some participants commented on the length of the NHS version stating that they felt it was too long.

I was mindful to provide adequate opportunity for each person to respond to the questions posed to the group. I made a particular effort to include those who were less talkative than members that were more willing than others to contribute. However, as the group conversation progressed all members were very willing to share their experiences and this was my observation across all five groups.

I felt that incorporating some of the techniques to successful interviewing that I have learned as a result of my counselling training was invaluable. In particular, the use of open questions and reflecting back key words or phrases proved useful in encouraging participants to elaborate. This was evidenced in the transcriptions and a review of the audio data. As identified by Hansen (2006) and often reported by other qualitative researchers it is common during the familiarisation and analysis phase of research to identify moments during the interview when perhaps you would have done things differently. For example, probed a little more or less, lengthened the duration of a pause or phrased a reflective statement differently. I too found myself identifying these moments on occasion. That said, across all the groups the conversation demonstrated good flow between topics and there was not any instances

when participants were struggling to comment on any questions that were posed to them. The majority of the conversation was generated from the participants and I felt that my posing the questions I asked them in the manner of a 'curious friend', asking for clarification when needed and encouraging elaboration, further facilitated the rapport, energy and flow of the conversations.

2.5 Discussion

The primary purpose of this phase of research was to generate a rich level of data that would be suitable to construct items for an exercise referral sensitive QoL scale. Twenty-three participants were interviewed across a total of five focus groups. Previous focus group investigations using exercise referral populations to explore mental health have utilised two groups in total (Crone-Grant & Smith, 1999), and in later studies four groups (Crone-Grant & Smith, 2001) and four groups with individual interviews (Crone et al., 2005). However, Crone-Grant & Smith's (2001) and Crone et al's (2005) study involved two pre-exercise referral and two post-referral focus groups which totalled the four groups. Similarly, the benefits of an exercise programme for individuals diagnosed with schizophrenia used a single group (Fogarty & Happell, 2005). One of the strengths of the current study is that five focus groups were conducted in total until the point of data saturation (Calder, 1977; Holloway, 1997) which enhances the likelihood that the experiences of exercise referral populations have been captured thoroughly.

While it is unlikely that a single scale will be relevant for each and every person who is referred for exercise, the age range of participants included in the current research (35-74 years) and the health conditions with which participants were referred are representative of individuals who typically attend exercise referral schemes. As a consequence this enhances the acceptability of the new scale. Those who had completed their initial 12-week period of exercise were purposively sampled. It was considered that these individuals had experienced an adequate exposure to exercise to enable them to comment upon their experiences of exercise-related life-quality. Selecting participants that had been newly referred ran the risk of generating less rich and informed data. However, it is also recognised that there is value in including those individuals who had less exposure to exercise, and perhaps in particular those who had failed to complete their initial 12-week referral. This population is currently underrepresented in the literature and there is a valued opportunity to learn more

about those reasons that prevent individuals from maintaining exercise behaviour from this population. Furthermore, a closer qualitative examination of the data generated for the current study may well give rise to a greater understanding of the relative importance placed upon codes/themes for males and females and across age groups. Although, an analysis of the results generated from responses to the final scale would also contribute to this understanding.

Integrating a pilot focus group into the study design provided a valued opportunity to assess the appropriateness of the questioning route for generating data sufficient to meet the research aims. In addition to additional probes that were introduced into the pilot focus group, the questioning route proved to be a useful tool to help guide the conversations in a focussed yet exploratory manner. It is likely that the period of iterative peer de-briefing that was undertaken during the development of the questioning route contributed to the overall suitability of the questions asked. Participants also responded well to the protocol of introductory, transition, key and ending questions (Krueger & Casey, 2000) that was utilised and as a result the questioning route remained unchanged for the subsequent focus groups. The pilot focus group also provided a valuable opportunity for reflection regarding the primary investigator's conduct and facilitation style although this reflective practice continued throughout the complete focus group data collection process.

A focus group approach to data generation and a two-phase approach to analysis (i.e., open coding and higher order thematic analysis) was particularly suited to meeting the research aim. Specifically, the participant demographics coupled with the methods employed for data generation and analysis in the current study facilitated an adequate depth and breadth of measurement for the scale under construction. The open coding level of analysis supports the construction of a desirable number of test items because of the volume and specificity of data (e.g., Streiner & Norman, 2008). In addition, the higher order themes informed the development of a conceptual framework of exercise-related life-quality. This conceptual framework contributes to ensuring that all four dimensions (i.e., leisure facilities and lifestyle physical activity, confidence and motivation, discomfort and fitness and mental and physical well-being) that participants identified as relevant to perceptions of exercise-related life-quality are adequately addressed by the measure under construction.

The results of the current study are unique in that they provide specific insights into what aspects of exercise-related life-quality are important at three distinct stages of exercise referral: at the beginning of a referral, during a referral and upon completion. The data has also highlighted what aspects of exercise referral are important at more than one stage. Having a stage-matched understanding of what factors may contribute to enhancing life-quality may assist practitioners to focus on those psychological, physiological, social and environmental aspects of exercise referral that are particularly responsible for positive changes in psychological outcomes at each stage. Stage matched models of this type are not unfamiliar to the psychology literature (e.g., Prochaska, DiClemente, & Norcross, 1992). However, because the primary outcome focus of the current study was to determine how exercise referral had impacted upon perceptions of life-quality regardless of the stage of referral, further work would need to be undertaken to establish what level of support could be attributed to these initial insights. If additional focus groups or individual interviews were used for example, this would require developing interview guides specifically designed for this purpose.

In addition to this unique contribution, some of the reports in the current study that describe aspects of exercise-related life-quality offer support for previous work involving exercise referral participants. However, published qualitative work for this population is somewhat limited. For example, to the authors knowledge Crone-Grant & Smith, (1999, 2001) represent two key conference communications and Crone et al. (2005) have been responsible for the only previous research that adopts a qualitative approach (specifically focus groups) to exploring the views of exercise referral attendees without targeting older women using individual interviews (e.g., Hardcastle & Taylor, 2001) or PA and well-being of older men and women using either group or individual interviews (Stathi, McKenna & Fox, 2003). For these reasons, some of the key findings between the current research and previously published work that has the greatest relevance in terms of the methods and participants used will be discussed below.

Crone-Grant & Smith, (1999) have reported similar findings to the current study by using a focus group approach to explore the broader mental health outcomes of exercise referral. In particular, data openly coded in the current study as comparisons to others while exercising, worry regarding exercising in an environment with others,

worry regarding participating in structured exercise and worry regarding the prospect of exercising in a fitness suite/gym have been described in similar terms. Crone-Grant & Smith (1999) described these concepts as 'apprehensions and anticipation' regarding starting a programme of exercise. These apprehensions and anticipations included feelings of 'self-consciousness' in the environment and 'lack of confidence concerning body image.'

The current study has been able to expand upon these reports by identifying that these apprehensions and worries actually reduced once participants had an opportunity to experience the leisure environment. A suggestion for future work would be to investigate whether a lack of reduction in such apprehensions contributes to the reasons why individuals fail to continue to attend. Indeed, there is significant potential for a focus group approach to explore the reasons for exercise referral attrition although it is recognised that this population may be potentially difficult to access by the very nature of their disengagement.

In the same study, the role of and contact with the instructor was identified as being important to exercise referral participants during focus group post programme reflections (Crone-Grant & Smith, 1999) and in a later study, Crone et al. (2005) identified once more that the professional and personal attributes of the instructor were important to the participant referral experience. The present study has also identified very similar reports but more specifically that the role of, and contact with, the instructor is of particular importance to exercise referral participants throughout their referral experience and beyond their initial 12-week period. Furthermore, participants in the present study were able to describe that the role of the instructor may be most valuable in terms of their knowledge and support level, in addition to communication style and a sense of familiarity. In this respect, the current research supports previous work but also extends the specificity of understanding regarding some of the psychological processes involved in the role and contact with the instructor.

Some of the reports made by participants in the current study support those reported in a previous qualitative investigation regarding the psycho-social aspects of physical activity and mental health (Crone et al., 2005). Crone and colleagues (2005) identified three higher order themes during their grounded theory driven work with

exercise referral participants. Crone et al. (2005) reported that the higher order theme described as 'social context' comprised social support, social network, and culture and environment. The theme 'actions' comprised playing a role, coping mechanisms and act of coping and the final theme 'outcomes' comprised sense of belonging, a sense of purpose and physical health.

Crone et al. (2005) identified that sources of social support were provided by scheme staff, other participants and family. Structured opportunities for exercise, social interaction and the importance of the time of day participants visited the leisure centre, particularly in terms of the type of exerciser they were likely to share their experiences with was important to participants. Similar to the reports provided by participants included in the current study, those included in Crone et al.'s (2005) study reported feelings of apprehension about attending the leisure centre and the levels of attention that are given by fellow exercisers within the leisure environment (openly coded as comparisons to others in the current study). Challenges using gym equipment and the importance of the role of the gym instructor in reducing these challenges were reports also shared with those generated from the current study. Crone et al. (2005) reported that the act of coping represented the belief that involvement in the exercise scheme assisted participants in coping with stressful life-events and thereby improved people's sense of well-being. Improvements in stress management were also reported by participants in the current study and it is entirely possible that perceptions in stress level were linked to improved well-being. An analysis of such relationships was beyond the scope of the current data and as such generates scope for future analysis of the existing data set.

Shared data reports between the current study and Crone et al. (2005) described how the number of staff, their professionalism, availability and personal qualities impacted upon the participant's sense of belonging and that such attributes were also critical to the referral experience and the likelihood that participants would complete their initial 12 weeks of referral.

Crone et al. (2005) referred to the importance of self-acceptance, recognising that social support, social norms and social interaction have been reported within the qualitative literature to support the mental health and PA relationship (e.g., Hardcastle & Taylor, 2001; Stathi et al., 2002). In addition to supporting these previous qualitative frameworks of the mental health and PA relationship, the results

of the current study can also add ‘comfortable with self’ and ‘self-awareness’ as related concepts to those previously reported.

Perceived barriers to physical activity have been identified previously and included physical, emotional, motivational, time, and availability (Biddle & Mutrie, 2001). Participants in the current study identified transport, time, childcare, poor health and the weather as the most significant factors to challenge a physically active lifestyle. An awareness of these typical barriers may help to support people through their exercise behaviour change and facilitate adherence through the recognition of potential challenges to PA, coupled with the development of individually tailored barrier reducing management strategies. Indeed, the depth of psychosocially focussed data generated by Crone et al (2005) led the authors to suggest that changes in training for exercise referral instructors is required to help support a shift from an over-reliance on the physiological processes and outcomes for those referred. This has clear implications for the purpose of the current thesis. Furthermore, with the limitations of the current study outlined earlier in mind it is noteworthy that participants identified that the role of the instructor was perceived as important at all stages of the referral programme. This makes Crone et al.’s (2005) suggestion particularly poignant.

Reported psychological outcomes and the statements made by participants regarding their motivations for continuing to exercise suggest that the desire to exercise continues in the absence of perceived significant physiological change. For example, one participant indicated, “I may not look any better but I feel better when I come out. You’re taller, you’ve got a spring in your step.” The reports made by participants regarding their association with perceived self-esteem and feeling better generally may help to assist with understanding what specific processes are involved when practitioners are often met with the ‘I just feel so much better’ response from participants who are assessed at programme completion. In this way, further exploration of the ‘feeling better – self-esteem’ relationship should be encouraged and this study has supported the justification for further qualitative research in this area. We know that psychological parameters are more likely to change during the typical course of an exercise referral scheme (DH, 2002). However, the measurement of psychological parameters tends to be neglected (e.g., Dugdill et al., 2005; Hardcastle & Taylor, 2001).

Coupled with a small but important number of previous focus group publications exploring the PA and psychological health of exercise referral participants (e.g., Crone-Grant & Smith 1999, 2001; Crone et al., 2005). The current study has emphasised that those involved in the 'grass-roots' delivery of exercise referral schemes should be encouraged and supported to collect valuable psychologically focussed data with the use of appropriate training and measurement tools wherever possible. The level of support offered from the findings of the current research and one of the most closely matched previous studies (i.e., Crone et al., 2005) would suggest that there is a relationship between psychosocial constructs of PA, mental health and exercise-related life quality and as such the extent of this relationship warrants future investigation. The current study has demonstrated that a focus group approach may well be suitable in meeting this aim. However, in terms of routine practice as yet, there is no measure of exercise-related life-quality available to help capture some of the meaningful outputs that have been highlighted here.

2.5.1 Limitations of the Research

It is acknowledged that there are some limitations of the current research. Participants who attended each of the five focus groups were recruited from a single referral scheme based within Nottingham, UK. The reasoning for this was logistical as the primary researcher had good links with this particular scheme which facilitated the engagement of staff to assist with the participant recruitment process. However, it is recognised that there would have been value in broadening the sample to include participants who attended other schemes across the country. Previous qualitative studies have utilised participants from a single scheme (e.g., Crone-Grant & Smith, 1999; 2001), although future focus group research may well benefit from including participants from a range of exercise referral schemes. This would contribute to the representativeness of the sample and may provide insights into whether differing scheme designs impact upon participant experience and in what way. As previously mentioned, it is also recognised that the participants included in the current study were predominantly female and is a sample characteristic shared by previous studies (e.g., Crone et al., 2005; Harrison et al., 2005; Lee, Griffin, & Simmons, 2009). However, there is scope for future research to reduce this disparity to allow for cross-comparison with the findings of the current and previous research.

2.6 Conclusions

A focus group approach to item generation proved to be an effective method of generating rich data suitable for the development of items for a test measure. A complete iterative/thematic approach to analysis which included both an open coding and a higher order thematic level of analysis was particularly appropriate to help generate as many items as possible but that would be relevant for subsequent studies of item generation and cognitive pre-testing (Chapter III). Higher order thematic analysis that included considerations of appropriate wording for items and whether the data represented causal or effect indicators of life-quality was a particularly useful stage of analysis to help clean the data and formulate a conceptual framework of exercise-related life-quality that would also inform the later phases of research that were undertaken as part of the current research programme. The results of the current study also support and extend previous qualitative investigations that have sought to explore the experiences and outcomes of exercise referral from a psycho-social perspective or with respect to mental health.

The following chapter reports the two initial phases of scale construction that were undertaken. Phase one used the focus group data to inform the design of items that were then cognitively pre-tested during phase two.

CHAPTER III Scale Construction:

Phases One and Two

3.1 Introduction: Item and Questionnaire Design

The previous phase of research (Chapter II), involved conducting focus groups (N = 5) to generate rich data suitable for item generation. With respect to item generation Streiner and Norman (2008) have identified that focus groups are a particularly beneficial approach. The volume and detail of data produced by employing a focus group method and the open coding and higher-thematic approach to the analysis will become particularly relevant during the following initial two phases of scale construction at the item level. What follows is a critique of the importance of the methods employed to questionnaire design both at the item level and in terms of the complete formatting and presentation of a test measure. These considerations were integral to the two phases of research undertaken for the initial construction of the scale.

Questionnaires are considered important research instruments, a tool for data collection with the primary function of measurement (Oppenheim, 1992). Employing a questionnaire to collect accurate data means writing items (questions) that enable respondents to provide the most accurate responses (Brace, 2004). A poorly written and/or structured questionnaire is likely to generate incomplete or inaccurate data and eliminating any items that may be ambiguous or incomprehensible at this stage of scale development has been identified as ensuring the interpretability of the measure (Streiner & Norman, 2008). Several texts have been published to guide practitioners through the crucial stages of questionnaire design and question formulation (e.g., Brace, 2004; Foddy, 1993; Hauge, 1993; Oppenheim, 1992) although research publications reporting the procedures undertaken for scale development generally include very little detail regarding this stage in favour of reporting their respective psychometric properties. Literature that helped to inform the construction of items in the current study are presented below.

Guidance regarding how to improve question wording for health behaviour surveys, particularly when aimed at culturally diverse populations has been offered. Warnecke, Johnson, chavez, Sudman, O'Rourke & Lacey et al. (1997) have argued that in the case of interview-administered measures, the investigator should envision the interview as a structured conversation in which ordinary conversational norms apply. Therefore, questions that might request redundant information or that are

threatening to the respondent should be asked in ways that minimize these effects. The authors also proposed attending to the wording and order of questions to ensure there is no question redundancy or respondent bias. Testing questions for potential racial or ethnic bias before using them was also recommended.

Similarly to the guidance offered by Warnecke and colleagues (1997), O'Brien, Ram, Vamvakas, and Goldman (2007) recently reported recommendations for changes to the Canadian Blood Donor Health Assessment Questionnaire (DHAQ). However, the recommendations made are transferable to scale development in general. By drawing upon learning from cognitive psychology, the authors outlined a series of cognitive processes that a respondent must undergo when completing a questionnaire namely: comprehending the question, retrieval of information and memories, judgement of the completeness and relevance of the information sought and finally mapping their response onto an appropriate category for example, yes or no. Using these cognitive proposals, O'Brien et al. (2007) proposed several recommendations for simplifying the DHAQ including: the use of capture questions (e.g., using the term Africa in place of listing all African countries), arranging questions into a logical order, selecting methods of administration that can elicit sensitive information (e.g., being mindful of socially desirable responses to questions during interview administration), reducing similar questions, simplifying questions, and testing questions and questionnaires before implementation.

Most recently Leitz (2010) has offered a comprehensive summary of the questionnaire design research literature and has proposed what the author argues is evidence-based and best practice guidance for questionnaire developers. Specifically, Leitz asserted that questions should be constructed to be as clear, simple, specific and relevant for the study's research aims as possible and should focus on current attitudes and very recent behaviour. The author recommends that more general questions should precede more specific questions and that vague quantifiers such as 'frequently', 'usually' and 'regularly' should be avoided. With respect to scaling measures and responses it was proposed that a desirable Likert-type response scale length ranges from five to eight response options. Leitz (2010) also argues that the inclusion of a middle option increases the validity and reliability of a response scale slightly because respondents have the option of providing a more neutral answer. Without a middle option, respondents may have to commit to endorsing an opinion that over or under represents their desired response. Furthermore, Leitz (2010)

proposed that the numerical scale should be uni-polar with matching verbal labels as anchors at both ends of the scale. The author offers examples of 'extremely' and 'not at all' as most effective verbal intensifiers and suggested that all numeric labels should be shown to respondents.

Rating scales or scaling measures are widely used by questionnaire writers (Brace, 2004). Although rating scales appear in several formats, all are designed to allow respondents to indicate the strength of their perception/relevance towards a specific topic (Foddy, 1993). They provide a straightforward way of asking information that is easy and versatile to analyse, and that provides comparability across time (Brace, 2004). The Likert scale, named after the scale's developer Rensis Likert is one of the most commonly used. Researchers developing Likert-type scales need to be mindful of both the question wording in terms of the guidelines suggested above but also the wording of the responses themselves and the most suitable number of response categories. For example, Lozano, Garcia-Cueto and Muniz (2008) have argued that between four and seven response options are desirable to optimise reliability and validity of a scale. While Skevington and Tucker (1999) have published a comprehensive guide to designing response scales for cross-cultural use in health care including recommendations for the wording of responses. The authors identified five different types of response scale which included intensity, capacity, frequency, evaluation and importance. Intensity responses are concerned with the extent to which a person experiences a state or situation, capacity refer to feelings or behaviour, frequency refers to the commonness of an occurrence, evaluation refers to state, capacity or behaviours that may require cognitive or affective appraisal and importance refers to the importance of each aspect of QoL for the respondent.

In summary, question wording, order, ambiguity, respondent options and burden and the overall presentation of the complete test measure (Brace, 2004; Foddy, 1993; Hauge, 1993) have been flagged as important for questionnaire and scale development. Additionally, the importance of pre-testing questions has also been raised (e.g., O'Brien et al., 2007; Warnecke, et al., 1997) which was addressed in the second phase of scale construction. For scale developers who utilise the Likert response, a recommendation of between four and seven response options has been given to increase the reliability and validity of a scale (Lozano, et al., 2008) and recommendations for the wording of scale responses has also been offered

(Skevington & Tucker, 1999). Save for cognitive pre-testing (phase two), all of these recommendations were given careful consideration during the questionnaire and scale response design phase of the current research and are documented more thoroughly in the Methods section (see 3.3).

3.2 Aims

The primary aims of phase one of the current research were to: a) construct data generated from focus groups into items that were suitable for new scale development, b) construct suitable response scales to the items that were devised and c) format and present these items and response scales into a complete test measure.

3.3 Methods

The open coding level of analysis used in the previous focus group study (Chapter II) helped to generate a large volume of items which is favourable during the early stages of item construction as some questions will not perform either conceptually or statistically well enough to be retained. The higher order thematic analysis informed a conceptual framework of exercise-related life-quality which ensured that each of the four domains reported by focus group participants (i.e., leisure facilities and lifestyle physical activity, confidence and motivation, discomfort and fitness and mental and physical well-being) were adequately represented by the test item pool (N = 50).

The inclusion of items that would represent each construct had already been considered in terms of whether such a question would represent either a causal or effect indicator of exercise-related life-quality during the latter stages of focus group analysis (Chapter II). This has been deemed as an important phase of QoL scale construction although distinguishing between causal and effect indicators can prove challenging as the two concepts are not always clearly distinguishable from one another (Fayers et al., 1997). As a consequence, this important step was re-visited ensuring that only those items considered as causal and/or effect indicators of life-quality were included in the item pool. Again, iterative phases of peer de-briefing (Spall & Stephen, 1998) were introduced into the research design to support this aim. It was during these peer de-briefing phases that a considerable amount of attention

was given to question length, wording and intended interpretability. Item construction was based upon the recommendations of Brace (2004), Foddy (1993) and Hauge (1993) and items were revised during the peer de-briefing phases until a consensus regarding the most appropriate construction of each of the 50 test items was reached. Some of the more specific considerations undertaken during item construction are detailed below.

Based upon general rules and guidelines developed by Belson (1981) and Oppenheim (1992), particular attention was given to question length, ensuring that questions did not contain sentences of more than twenty words. 'Double-barrelled' questions (i.e., questions which would require more than one response) were avoided and although 'don't know' or 'not applicable' response options were not included, the response scales were designed in such a way that a 'not at all' response, for example, allowed a respondent to indicate the lowest level of importance of a particular question to their perceived exercise-related QoL status. Similarly, the response scale 'never' was included to include a zero level of frequency. Attention was given to avoiding acronyms, abbreviations, jargon and technical terms. Due consideration to memory recall was given by requesting that respondents consider their lives in the context of the questions asked no longer than during the last two weeks (Skevington et al., 2004). Attention was also given to: a) the objectives of the questionnaire, b) how the measure was likely to be administered, c) the knowledge and interests of the respondents to be targeted, d) the introduction (instructions) to the questionnaire, e) the order of questions, f) the type of questions, g) the possible responses, and finally h) how the data would be processed (Hague, 1993). The introduction (instructions) to the test items was considered using that from a previously validated generic QoL measure (Skevington et al., 2004) and question order was considered in terms of how items were grouped together into topics which followed a logical sequence (Hague, 1993).

Responses to questions were developed as a 5-point Likert scale and based upon the recommendations of Skevington and Tucker (1999). The structure and presentation of the questions and responses of the WHOQOL-BREF (Skevington et al., 2004), which was added for the purposes of convergent validity analysis, also adopts the response scales reported by Skevington and Tucker (1999). Consequently, the WHOQOL-BREF was used as a template for both the presentation and formatting of

test questions and the accompanying Subjective Exercise Experience Scale (SEES; McAuley & Courneya, 1994) which was also added to the 50 test questions for the purposes of later phases of scale construction - convergent validity analyses (see Chapter V). This procedure ensured formatting consistency across the complete test questionnaire. The rationale for the selection of the WHOQOL-BREF and SEES as construct validity measures is discussed in more detail in Chapter V. However, it is worth noting the addition of these measures here for the purposes of understanding the overall appearance and formatting of the complete test measure that was administered. The total number of items comprised 50 test items, 12 SEES items and 26 WHOQOL-BREF items resulting in 88 items in total.

3.4 Resultant Items and Scales

Fifty items were constructed and included in the draft test questionnaire of which nine were reverse-scored questions (i.e., a higher rating represents a poorer exercise-related life-quality perception). Table 9 details the wording of each item belonging to the construct leisure facilities and lifestyle physical activity, Table 10 details the wording of each item belonging to the construct confidence and motivation, Table 11 details the wording of items belonging to the construct discomfort and fitness and Table 12 details the wording of each item belonging to the construct mental and physical well-being. A general question regarding health status was not included as respondents are requested to identify the specific health reasons for which they have been referred in the demographic questions contained within the new scale. It is likely that this information alongside other more clinical information that is typically shared as part of the referral process would provide adequate insights into respondent health status at least from a physiological perspective. The aim of the new scale was to provide an opportunity for exercise referral practitioners and those referred into schemes to gain a more detailed understanding of perceptions of exercise-related life-quality. As a consequence, the items contained within the new scale were devised to avoid some of the difficulties discussed earlier with the notion of defining ‘health’ in more general terms.

Table 9. Questions reflecting the construct leisure facilities and lifestyle physical activity

Questions Reflecting the Construct Leisure Facilities and Lifestyle Physical Activity
How important to you is the affordability of leisure facilities in order to participate in a structured exercise programme?
How much do you feel you have a choice regarding the exercise you undertake?
How much do you feel you have a choice regarding the time at which you exercise?
*How much do you feel that lifestyle factors (e.g., transport, time, childcare, poor health & the weather) affect your ability to be physically active?
*How much do you feel you compare yourself to others while exercising?
*To what extent do you worry about exercising in an environment with others?
*To what extent do you worry about participating in structured exercise?
*To what extent do you worry about the prospect of exercising in a fitness suite/gym?
*How much does any injury you may have prevent you from being physically active?
How much do you currently enjoy physical activity?
How much is social contact with others a part of your current lifestyle?
How important is it for you to manage your weight?
How much do you feel that you incorporate physical activity into your daily lifestyle?
How well do you feel you adhere to eating habits that are beneficial to your health and any illness you may have?
In terms of exercise, how competitive are you?
In terms of exercise, how determined are you?
In general, how much opportunity do you have to make time for yourself?

Note. Items (n = 17) * denotes a reverse-scored item.

Table 10. Questions reflecting the construct confidence and motivation

Questions Reflecting the Construct Confidence and Motivation
How confident are you using gym equipment?
How confident are you that you can exercise in a leisure centre with minimum support?
How confident are you in your ability to participate in regular physical activity and exercise?
How confident are you in your ability to undertake daily tasks that do not necessarily involve physical activity?
In general, how confident are you around other people?
How confident are you in your own physical ability?
How confident are you to join in activities with family and friends?
How important is it to you right now to maintain a physically active lifestyle?
How confident are you, right now that you can maintain a physically active lifestyle?
How motivated are you to maintain a physically active lifestyle?
How likely are you to maintain a physically active lifestyle?

Note. Items (n = 11).

Table 11. Questions reflecting the construct Discomfort and fitness

Questions Reflecting the Construct Discomfort and Fitness
*How often do you experience physical pain?
*How often do you feel you have to manage any symptoms of pain?
*How often do you feel you have to manage the symptoms of any illness you have?
How often do you feel you have achieved something in terms of your physical fitness?

Note. Items (n = 4) * denotes a reverse-scored item

Table 12. Questions Reflecting the Construct: Mental and Physical Well-being

Questions Reflecting the Construct Mental and Physical Well-Being
How would you rate your energy levels?
How would you rate your outlook on life?
How would you rate your level of self-esteem?
How would you rate your physical health?
How would you rate your mental health?
How would you rate your sense of well-being?
How would you rate the quality of support you receive from leisure centre staff to be physically active?
How would you rate the quality of support you receive from family and friends to be physically active?
How would you rate the quality of your sleep?
How would you rate your memory function?
How would you rate your level of motivation to be physically active?
How would you rate your ability to manage stress?
How would you rate your ability to manage your mood?
How would you rate your ability to manage your weight?
How would you rate your current life expectancy?
How would you rate your ability to undertake everyday tasks that require some level of physical activity?
How would you rate your overall quality of life?
How would you rate your current knowledge of the benefits of physical activity and exercise for health?

Note. Items (n = 18).

The Likert response scale utilised three examples proposed by Skevington and Tucker (1999). In particular response scales were chosen to assess frequency (Figure 9), capacity (Figure 10) and evaluation perceptions (Figure 11).

Figure 9. Example of the response scale to assess frequency

	Never	**Seldom**	**Quite often**	**Very often**	**Always**
29) How often do you experience physical pain?	1	2	3	4	5

Figure 10. Example of the response scale to assess capacity

	Not at all	**Not much**	**Moderately**	**A great deal**	**Completely**
11) How much is social contact with others a part of your current lifestyle?	1	2	3	4	5

Figure 11. Example of the response scale to assess evaluation perceptions

	Very Poor	**Poor**	**Neither good nor poor**	**Good**	**Very good**
49) How would you rate your overall quality of life?	1	2	3	4	5

Once phase one of scale construction had been completed and the 50 test items and corresponding response scales had been developed, phase two commenced. This second phase of scale construction aimed to cognitively pre-test the items and scales to check that they were understood and interpreted as intended. Some relevant background to and critique of this method is presented in the following section.

3.5 Introduction: Cognitive Pre-Testing

A frequent difficulty with questionnaire design is that respondents commonly misinterpret questions and this difficulty has been recognised within the literature (e.g., Belson, 1981; Hunt, 1982; Nuckols, 1953). In some instances, scale developers may employ an expert panel approach to enhance the likelihood that questions will be understood and interpreted as they were intended. In this way, any amendments to items are made solely by a team of experienced researchers and/or practitioners. However, cognitively pre-testing newly developed scale items with a representative sample of individuals for whom the new scale is intended provides a valuable opportunity for in-depth field testing. Arguably, pre-testing items with a population in such a way allows for new items to be tested and piloted more thoroughly than an expert panel approach would allow because representative respondents are given the opportunity to feed directly into this critical phase of scale development.

Pre-testing is a method of checking that questions work as intended and are understood by those individuals who are likely to respond to them. It is also the case that pre-testing has the capacity to reduce sampling error and increase questionnaire response rates (De Leeuw, 2001; Drennan, 2003). The method also has the capacity to gain valuable feedback regarding the overall presentation, formatting and relevance of the measure being tested for the population for whom it is intended. Such considerations are often referred to as face validity (Rust & Golombok, 2009) or how practical and realistic the measure is to complete (Murphy & Davidshofer 2001) and as such are often taken for granted as being a common-sense approach to test development (Streiner & Norman, 2008; Murphy & Davidshofer, 2001). However, mistakes are easy to make more particularly perhaps because test developers can became overly familiar with medical and jargon language, for example that may find its way into item wording.

The way in which pre-testing should be undertaken (i.e., the pre-testing methodology) has been the source of much debate, particularly in terms of the lack of standardisation of the method. Few studies have been published that detail specific pre-testing protocols with an appropriate level of detail in of the methods employed or guidance for researchers (Presser et al., 2004). This includes the consistent recording of interviews (Blair & Presser, 1993) and an evaluation of the methods

used (Collins, 2003). Nevertheless, cognitive interviewing is often the method of choice for pre-testing questions and is considered an important part of the design research process - the only way to determine in advance whether a questionnaire causes problems for interviewers or respondents (Presser et al., 2004) and also as a valuable addition to psychometric techniques when validating complex tools (De Silva, Harpham, Tuan, Bartolini, Penny, & Huttly, 2006).

An example of one of the least structured yet effective approaches to cognitive interviewing has been provided by Chong, Chay and Chuen (2006). During the validation phases of a childhood asthma questionnaire for Asian children, the authors simply asked a group of young asthma sufferers who were purposively selected to represent a range of age, sex and asthma severity (N = 10) to complete two questionnaires validated in English while under observation. The respondents were encouraged to ask the observer whenever they had difficulties responding to any of the items. However, probes regarding the overall impression of the measure including difficulty in understanding the questions or instructions and the relevance for them were saved until the questionnaire had been completed. This phase of pre-testing played a critical role in the researcher's choice of the most appropriate childhood asthma measure to validate as a Singaporean version. For example, the authors report that the children could not really tell the difference between words such as 'troubled', 'frustrated' and 'angry' or between 'uncomfortable' and 'different or left out' and these considerations impacted directly upon the measure selected for translation.

Chong et al. (2006) did not make use of pre-planned probe testing protocols but rather adopted a much less structured approach to pre-testing. Instead the study design relied on researcher observation and post completion commentary from respondents. However, this relatively un-structured approach to cognitive pre-testing still helped Chong and colleagues (2006) to identify problems with question and response option wording which directly informed later phases of their research. In this respect, these study outputs reflect how even a relatively simple approach to pre-testing can give rise to such meaningful data.

Foddy (1996) has offered a more complete critical appraisal of methods and identified that is typical for the purpose of pre-testing that: a) respondents will be asked to think out loud while completing the test questionnaire and/or b) the

interviewer will introduce probe questions to check that the questions are understood and being interpreted as intended. This appraisal informed the approach to pre-testing adopted by the current study. Willis (2005) has published a comprehensive text that details various approaches to cognitive interviewing and this text was also used to support the selection of appropriate methods in the current study. Willis (2005) argued that utilising only the think-aloud technique is difficult and that probe questions tend to encourage think-aloud behaviour. Willis (2005) also recommends that a combination of both methods usually removes the need to provide special think-aloud instruction to participants that some find difficult. Consequently, cognitive interviewing is best characterised as a combination of think-aloud and probing procedures.

3.6 Aims

The aims of this study were to 1) ensure that the questions developed for the scale under construction were interpreted as they were intended, and that questions could be responded to with clarity and ease and 2) ensure appropriate face validity of the test measure.

3.7 Methods

In each case, pre-testing was facilitated by the primary investigator. Ethical approval was obtained from Nottingham Trent University and the UK National Research Ethics Service (NRES; reference number 06/Q2404/36; date of approval 25/4/2006). Participants for all 15 cognitive pre-tests were recruited through a local exercise referral scheme in Nottingham, UK. All participants were provided with an NRES template information sheet describing the study (Appendix 5). A shorter, quick reference version was produced that was specifically relevant for self-complete participants (Appendix 6) and one for interview and telephone administration (Appendix 7). All participants provided full written consent to their participation, again using the NRES consent form template (Appendix 3).

Fifty questions were subjected to cognitive pre-testing protocols that were designed based upon the recommendations of Willis (2005). It is usual that interviews are conducted with five to ten people (Willis, 2005). However, a greater total number of

participants were included in the pre-testing phase because the questionnaire was administered in three different ways, namely: a) self-complete with think-aloud and question probes (n = 5), b) interview-administered with think-aloud and question probes (n = 5) and c) telephone administered with think-aloud and question probes (n = 5).

An initial testing protocol was developed prior to pre-testing using the recommendations of Willis (2005; Appendix 8) and each of the 50 questions that were subjected to pre-testing were allocated corresponding probe questions that reflected areas of clarification as appropriate. For some questions, these probes were quite specific. For example, to determine if the terms 'physical activity' and 'exercise' in the same question would cause confusion, participants were asked, "the question uses the words physical activity and exercise in the same question. Does that sound OK to you or would you use something different?" Other probes were more general and included questions such as "how did you arrive at that answer?", "was that easy or hard to answer?", "I noticed that you hesitated, tell me what you were thinking." The conduct of the primary investigator, including the delivery style of the questions and probes was similar to that employed for the focus group phase of research (Chapter II). The intention was that each interview conducted was collaborative in nature. Participants were encouraged to generate the majority of the conversation, while the primary investigator introduced both the written and additional probes at key points throughout the interview. The primary investigator was able to draw upon experience and training in counselling methods such as MI and Solution Focussed Therapy (SFT) to build rapport during the interview, facilitate the discussions and to elicit as much information as possible from respondents in terms of recommendations for item amendments.

As pre-testing progressed and greater clarity was achieved regarding the interpretation of the questions by respondents, the initial probe protocol was amended slightly on two further occasions (Appendix 9 and 10). For example, the question "how much do you feel that you incorporate physical activity into your daily lifestyle?" Had probes: "what does the term physical activity mean to you? In what way does this differ to the word exercise, if at all? These were included in the initial protocol. However an additional probe was introduced into a second version of the probe protocol: "what sorts of things come to mind when you think of incorporating physical activity into your daily lifestyle?" This allowed for an

exploration of the term physical activity in addition to prompting for views regarding what is meant by daily lifestyle physical activity. Similarly, an example of an amendment that was made to the probe protocol on the third and final occasion was in response to the question "how much do you feel that lifestyle factors (e.g., transport, time, childcare, poor health & the weather) affect your ability to be physically active?" The initial and second probe protocol asked: "how easy or hard was it to choose an answer?" The third protocol also included an additional probe: "you will notice that I provided you with some examples of lifestyle factors (transport, time, childcare, poor health and the weather) did these examples help you, or make it more difficult to answer the question?" This allowed for a greater exploration regarding whether the inclusion of differing examples within one question were viewed as problematic in addition to a more general understanding of item response difficulty.

Based on the recommendations of Willis (2005) as the interviews progressed, additional probes were introduced to extend the primary investigator's understanding of the interpretation of the test questions (Appendices 10 and 11). For the purposes of face validity considerations, participants were encouraged to comment on the complete test measure including formatting, presentation and relevance at the end of the interview. Each interview was digitally recorded and notes were taken throughout. The interviews were listened to on a minimum of two occasions and notes made during the interview process were combined with any additional notes made from retrospective reviews of the audio data.

3.7.1 Data Handling and Analysis

According to Willis' (2005) recommendations, the pre-test data set was subjected to procedures as follows:

1) Cognitive interviewing outcome reports that summarized the results of each of the three conditions under which the questionnaire was administered were produced.

2) Summary data records were given: a) qualitative consideration of what the problems were and whether they were similar across interviews, and b) quantitative

consideration of the frequency with which problems emerged, to gain insights into the severity of the problem.

3) The complete pool of participant responses and recommendations underwent an iterative phase of peer-debriefing whereby the data were critically reviewed by the research team and a consensus regarding question wording was reached.

The cognitive interviewing outcome reports (Appendix 11, 12 and 13) were generated from carefully reviewing the audio data and the accompanying notes made by the primary investigator for each interview conducted under each of the three conditions (self complete, interview and telephone). For each question, the reports documented whether if any problems were experienced by participants in responding and if so, the nature of the difficulty. These summary reports revealed both the frequency and nature of item difficulties across all three administration methods and were used to generate an overview of item performance (Table 14). The resultant table and individual summary reports were reviewed by the research team and consideration was given to each of the items that had been identified as problematic in terms of whether these items should be amended or removed from the test-item pool.

3.7.2 Participants

Twelve females and three males (N = 15) were recruited for the purposes of pre-testing and this cohort of participants was exclusive to the cognitive pre-testing phase of research. Participants ranged in age from 36 – 76 years (M = 60, SD ± 10 years). Table 13 indicates at what stage of referral the participants were at when the pre-testing was conducted. The employment status of participants included employed (n = 4), retired (n = 10), and unemployed (n = 1). The reasons for referral included weight loss, asthma, diabetes, hypertension, depression, mobility and joint difficulties, smoking cessation and post operative and cardiac rehabilitation (Table 13). These demographics are typical of those who are referred into 12-week exercise programmes.

Table 13. Pre-testing participants demographic data.

Gender	Age	Test Method	Referral Reason	Referral Stage
Female	48	Self-Complete	Weight Loss	4 weeks
Female	64	Self-Complete	Weight Loss, Hypertension & Hip Replacement	2 weeks
Female	58	Self-Complete	Post Operative Exercise	6 weeks
Female	60	Self-Complete	Hypertension & weight loss	Complete
Male	62	Self-Complete	Post Operative exercise (hip replacement)	Complete
Female	68	Interview	Depression & Borderline hypertension	2 weeks
Female	65	Interview	Arthritis & weight loss	12 weeks
Male	40	Interview	Cardiac rehabilitation	Complete
Female	76	Interview	Weight loss, smoking Cessation, Osteoporosis & Asthma	11 weeks
Male	61	Interview	Cardiac rehabilitation	Complete
Female	63	Telephone	Asthma	6 weeks
Female	36	Telephone	Mobility difficulties, Diabetes & depression	Complete
Female	68	Telephone	Weight loss	2 weeks
Female	69	Telephone	Hypertension & Diabetes	Complete
Female	64	Telephone	Depression	6 weeks

3.8 Results

In light of Willis' (2005) previously cited recommendations, qualitative and quantitative consideration of the frequency with which problems emerged gave rise to the following amendments. The wording of six questions was changed to reflect the recommendations of respondents, one question was split into two separate questions for clarity and accuracy and one question was removed completely because respondents considered that is was too general (Table 14). There were no reported difficulties in understanding the response options or their suitability to their respective questions.

Table 14. Summary of pre-testing results: Self-complete, interview and telephone administered

Question	Results & Source of Problem	Response
How much do you feel you have a choice regarding the exercise you undertake?	Choice of exercises during exercise classes vs. mode of exercise needs clarification (1 x self complete)	Amend question to include 'type' of exercise
How much do you feel you compare yourself to others while exercising?	No difficulties understanding the question or response. Two suggestions to remove "you feel" (1 x self complete 1 x interview)	Amend question to remove 'you feel'
How competitive are you?	Four suggestions to clarify the context (2 x interview 2 x telephone)	Amend question to include 'in terms of exercise'
How determined are you?	Four suggestions to clarify the context (2 x interview 2 x telephone)	Amend question to include 'in terms of exercise'
How often do you feel you have to manage the symptoms of any illness you have, if any?	Four suggestions to remove the term "if any" (3 x self complete 1 x interview)	Amend question to remove 'if any'
In general, how would you rate your level of motivation?	Three suggestions for greater clarity regarding if in general or with specific reference to physical activity (2 x interview 1 x telephone)	Remove question.
How would you rate the quality of support you receive from others to be physically active	Clarity required to define 'others' (2 x self complete)	Question split in two. One for leisure staff, one for family/ friends
How much do you worry about the prospect of exercising in a gym	The term 'fitness suite' is often used to refer to gym facilities (interviewer observation & additional probes)	Amend question fitness suite/gym

Cognitive interviewing outcome reports for each of the three methods that were tested can be found in Appendices 12, 13 and 14. The reports detail more general comments from the questioning probes used in addition to those that resulted in the amendments outlined in Table 14. What follows is a detailed account of the feedback provided by participants (N = 15) across the three cognitive pre-testing methods including pre-testing results that did not give rise to amendments.

The introductory instructions describing the purpose and method of completion of the questionnaire posed no difficulty with respect to understanding or interpretation. Comments provided by respondents indicated that the instructions were "clear", "very clear" and that there was "no difficulty at all" in understanding what was being requested.

Questioning probes were designed to test respondents' understanding of the term 'structured exercise'. Pre-testing revealed that respondents interpreted the term as referring to exercise that was "supervised" or "organised", "exercise with a leader" or that was "undertaken at a particular time". These interpretations of 'structured exercise' were as the primary investigator intended.

One of the questions tested included examples of lifestyle factors that the focus group participants had identified as being potential barriers to exercise participation. These examples were transport, time, childcare, poor health and the weather. Probes were developed that aimed to clarify if including these examples in the question helped or hindered a response. Generally, it was felt that the examples helped respondents to complete the question. Indeed for older people, 'childcare' allowed respondents to consider commitments to the care of grandchildren. When participants were asked if removing the examples altogether would make the question clearer, only one respondent agreed. Similarly when questioned as to whether providing a separate question for each example would add clarity to interpretation, a single respondent agreed, but acknowledged the potential increase in respondent burden due to the increase in questionnaire length by employing this amendment. For these reasons, the question remained unchanged.

Questioning probes were developed to establish what respondents understood by the terms ‘physical activity’ and ‘exercise’ and if including both terms in the same question was problematic. Results from cognitive interviewing across all three methods of administration revealed that including both terms within the same question posed no difficulties with the understanding, interpretation or the ability to respond to these questions.

When questioned if any injury prevented the respondent from being physically active, probes were introduced to determine if the response scale would account for respondents who did not consider themselves to have any injury. This was particularly important to explore as none of the response scales devised by Skevington and Tucker (1999) allow for a ‘not relevant’ option. The question asked “how much does any injury you may have prevent you from being physically active?” In this case the response options were ‘not at all’, ‘not much’, ‘moderately’, ‘a great deal’ or ‘completely’. Cognitive pre-testing determined that the response ‘not at all’ was suitable and selected by those who had no injuries to report.

The test questions “how competitive are you” and “how determined are you” were two of the most open questions included in the draft test measure and it was anticipated that these questions may subject to misinterpretation. General probes for these questions included: “how did you arrive at that answer?”, “was that easy or hard to answer?”, “I noticed that you hesitated, tell me what you were thinking”. Pre-testing revealed that some clarity regarding the context of competitiveness and determination would be required (i.e., generally or with respect to exercise). Two subsequent questioning probe protocols that were developed following initial testing included more specific probes that explored the level of specificity needed to respond to the question accurately (Appendices 10 and 11). Two respondents who had completed the questionnaire under interview conditions and two respondents who had completed under telephone administration conditions reported that the question would require amending to reflect the specificity of general competitiveness and determination or with respect to exercise behaviour. Interestingly, those respondents who self-completed the questionnaire did not report any such difficulties. A close inspection of the qualitative data used to generate the items indicated that amending these questions to refer to exercise behaviour was more appropriate.

Respondents were asked how confident they were to exercise in a leisure centre with minimum support. One respondent who self-completed the questionnaire reported that greater clarity may be required regarding the source of support; for example, from friends and family or from an exercise instructor. Because the frequency of the difficulty was so low (i.e., a single report) and because another test question that targeted perceived support from "others" was amended and split into two to identify: a) support from family and friends, and also b) an exercise instructor to be physically active, this 'leisure centre support' question remained unchanged.

Pre-testing indicated that the word 'adhere' included in the question "how well do you feel you adhere to eating habits that are beneficial to your health and any illness you may have?" may prove problematic. During an iterative phase of peer de-briefing (Spall & Stephen, 1998) the research team proposed alternative words and phrases such as "stick-to", "sustain" "maintain" and "uphold". It was decided that the question focus was regarding the maintenance of healthy eating habits and for this reason a consensus was reached for the question to be re-worded to "how well do you feel you maintain eating habits that are beneficial to your health and any illness you may have?"

In terms of face and content validity considerations, participant responses to ending probes that encouraged feedback on the ease or difficulty with which they completed the test and if they had any suggestions for further development/amendments to the measure indicated that the measure was easy to complete and relevant to them. Participants' reports included that they had "no difficulty at all" with completing the questionnaire and that it was "easy" to understand and complete. A 64 year old female participant stated that the questions were "particularly relevant to older people," explaining that from her own experience, some of the questions she had been asked throughout the course of her contact with medical professionals had been less relevant to her daily lifestyle than others. No suggestions were made for the future development or amendments to the measure.

The value of reflection as part of a qualitative research approach has been discussed in an earlier chapter (Chapter II). Phase two of the current research was largely qualitative in nature and so what follows in the next section are reflections of the

interviewer regarding the processes involved in undertaking this phase of research and the value of cognitive pre-testing as an approach to evaluating item rigour and suitability.

3.9 Reflections of the Interviewer

The complete process of undertaking cognitive pre-testing including the skills required to prepare for and conduct the interviews were very similar to those required for the focus group phase of research (Chapter II). I am sure that even simple preparation tasks such as ensuring that the probe protocol and enough copies of the test questionnaire had been printed the day before the interviews were to be conducted added to my feeling of 'readiness' to undertake the pre-testing. I found the participants only too willing to complete the test-measure despite it containing 50 items which I actually thought may have been quite burdensome. From my view this was true for each of the conditions under which the test items were administered and I felt that a rapport developed quite quickly.

Similar to the focus groups that I facilitated I was mindful to keep probing questions open and to pay minimum attention to the probing protocol questions in favour of engaging more with respondents. At times it was difficult to refer to the probe protocol while administering the test items under interview conditions as I had to read the test items aloud accurately and ensure that I introduced the correct corresponding probe whether this be general or specific. This became easier the more interview based pre-tests I conducted as I became more familiar with the probes associated with each item and therefore felt less reliant on the probe protocol.

Some of the notes I made during the interviews were helpful in terms of having an immediate sense of the frequency and severity of problems with items. In the main, questions were interpreted as intended, and therefore these notes were made quite infrequently which allowed for them to be quite detailed. I could also use these notes to review all the problems that had been reported once all 15 interviews had been conducted before I even reviewed the audio data. I think the data recording methods of note taking and audio recording were particularly useful as this ensured that no problems were missed. I also found the structure that was provided from adopting Willis' (2005) approach to pre-testing extremely valuable. I had a protocol to follow,

I knew what probes I would be introducing into each interview. The probe protocols were invaluable in this respect and some of the additional probes I introduced were useful in terms of clarifying any difficulties respondents may have had with responding to items.

At times it was difficult not to enter into purposeful conversations with respondents regarding their experiences of exercise referral. This was because much in the same manner as the focus groups, more often than not respondents were keen to provide some supporting commentary and context to the responses they were providing. This required quite a lot of discipline on my part to stick to the primary purpose of the interviews which was to test the performance of the items. As I reviewed the audio data and reflected on the enjoyable conversations I had with the respondents, I wondered what experiences could have been evoked from and reported by participants had the aims of this study been more exploratory. Certainly, some of the reports made by respondents were quite similar to those made by the focus group participants and it was difficult for me not to encourage elaboration in these instances.

I attribute the quality of the data to the willingness of the respondents included in this study to support the research process. I think I did quite well at reminding myself of the purpose of the interviews which was very different to the kind of qualitative interviews I am used to, perhaps more particularly having a counselling background it was tempting to probe much more on some of the respondent's commentary. In this respect, I think this mindfulness also contributed to the quality of the data produced. Again, the procedures proposed by Willis (2005) that were followed including the production of outcome reports were a valuable part of the research process and helped to clarify and quantify any difficulties with items in a manner that could be easily communicated between the research team. I genuinely enjoyed undertaking the interviews and I feel that in addition to improving the rigour of the test measure, the skills required to conduct them has added a new dimension to my professional development.

3.10 Discussion

The purpose of the current research was to 1) construct robust items and response scales alongside an appropriate presentation style suitable for an exercise referral sensitive measure of life-quality and 2) cognitively test the performance of these items and scales to ensure that they were understood and interpreted as intended and that the overall presentation was also appropriate to ensure face validity. These two distinct research aims were undertaken in two phases. What follows is a critique of how well the methods employed by each phase of scale construction met these aims with reference to relevant literature where appropriate.

The construction of the scale was guided by a number of key texts (e.g., Brace, 2004; Foddy, 1993; Hague, 1993; Oppenheim, 1992; Streiner & Norman, 2008) and overseen by the principles of brevity, simplicity and concreteness (Foddy, 1993). The instructions as to how to complete the test questionnaire, the response scales developed and general format were all based upon an existing validated measure of general life-quality (Skevington et al., 2004). Such attention to the detail of item construction in this manner aimed to ensured adequate interpretability of the measure (Streiner & Norman, 2008). However, cognitive pre-testing revealed problems with eight questions in total. The number of problems may have been minimised by the procedures undertaken during phase one of the current research but this highlights the value of undertaking cognitive pre-testing during the initial phases of scale construction at the item level. Had cognitive pre-testing not been undertaken, these problems would have been carried forward into the psychometric phases of research to follow.

A necessary requirement of a successful question-answer cycle is that both researcher and respondent have a shared understanding of the topic under investigation (Foddy, 1993). The researcher should also try to avoid putting respondents in the position of having to entirely invent or fabricate answers on the basis of respondents being provided with little information (Foddy, 1993). The importance of the psychological evaluation of exercise referral schemes including the relationship between exercise and QoL was reviewed thoroughly by the primary investigator prior to commencement of cognitive pre-testing. The primary investigator also had a good understanding of the operational aspects of the exercise

referral schemes that respondents were attending gained from previous employment. As a consequence the primary investigator was in a desirable position in terms of understanding the topic under investigation. Additionally, explaining the purpose of the study in an informal manner to respondents undoubtedly facilitated a shared understanding of the study focus and what was required from participants. In this respect these contributions to the overall study design addressed the challenges identified by Foddy (1993).

Careful attention was paid to ensuring that the appropriate methods had been matched to the aims of each of the two phases of the current research and ongoing reflection of the research process (Bennett-Levy, 2003) contributed to the construction of test items that required a fairly low level of amendments (i.e., 8 item amendments from a pool of 50) which is comparatively less than identified in examples of previous research. For example, Nuckols (1953) reported that one in six participants incorrectly re-defined a test question presented to them when asked to explain the question in their own words. Two items in the current study adopted this re-wording approach to item testing. These items were: "how confident are you in your ability to participate in regular physical activity and exercise? And "how would you rate your current knowledge of the benefits of physical activity and exercise for health?" Neither item posed mis-interpretation difficulties. More recently, Belson (1981) reported that only 29% of respondents offered the intended interpretation of a question and, moreover, that the highest score of accuracy for any of the questions tested was only 58%. It is important to be mindful that Belson (1981) chose to test those questions deemed to be particularly problematic from a review of existing measures and that are typical of those avoided during phase one of the current scale under construction (e.g., questions that required more than one answer). This may account for the particularly frequent incidences of misinterpretation reported by Belson (1981) and also reinforces the value of employing such a rigorous approach to phase one of item construction in the current study.

The methods employed by phase two of the current research (cognitive pre-testing) were particularly well matched to the research aims and represent an approach to pre-testing that is particularly comprehensive. For example, the constructed items were tested under three conditions: self-complete, interview and telephone administered. It is more usual that if scales under construction are pre-tested, only a single condition (typically interview-administered) is employed (e.g., Wildy & Clarke, 2009). Testing

the performance of items under three different conditions and the results generated supports the flexibility of administration and thus facilitates the use of the scale in practice.

Additionally, the combination of both think-aloud and probe techniques alongside evolved probe protocols that responded to the changes in the depth of understanding regarding question performance was particularly effective. Specifically, this approach to pre-testing allowed for flexibility within an otherwise structured design which complemented the overall rigour of the research in terms of how the data were collected and reported. Previous studies have explored the use of think-aloud technique only (e.g., French, Cook, Mclean, Williams, & Sutton, 2007). However, French et al. (2007) discuss the results of applying think-aloud techniques within the context of the performance of a Theory of Planned Behaviour questionnaire in comparatively more detail than a critique of the method used in general. In turn, this limits the learning to be gained from the experiences of applying the think-aloud technique only by reporting the results in such a manner. A further limitation, but that is recognised by French et al. (2007) was that utilising the think-aloud approach in isolation required participants to verbalise their thoughts and if this is not done effectively, problems remain undetected.

The current study employed both think-aloud and probe techniques and as a consequence this limitation was minimised. Similarly, a) providing detailed probe protocol reports regarding the specific methods used for each question tested, b) thorough reporting of the results and c) implementing an iterative phase of peer de-briefing (Spall & Stephen, 1998; Willis, 2005) has added to the transparency and rigour of the research. It is also expected that such a level of detail reported for the cognitive pre-testing phase of research in the current study may contribute to lessening the difficulties identified by Presser et al. (2004) regarding the poor level of detail reported of the methods used for pre-testing and lack of detail regarding specific pre-testing protocols. The complete two-phase process of scale construction employed by the current research may also assist in bridging the gap regarding guidance during the pre-testing phase of scale construction research which has also been identified by Presser et al. (2004). Furthermore, the results reported in the current study support what has previously been identified as face validity criterion (e.g., Murphy & Davidshofer 2001; Rust & Golombok, 2009).

The response options included in the test measure were constructed by Skevington and Tucker (1999). No problems were reported regarding the response options either in terms of understanding the wording or the appropriateness of the question to which they were allocated. The iterative phases of peer de-briefing undertaken during phase one of the current research undoubtedly contributed to the process of ensuring that the most appropriate response scales were matched with each item. However, the careful selection of response scale options appropriate to the design of the scale under construction, particularly those that have been grounded in scientific research and field tested (Skevington & Tucker, 1999) also reduced the likelihood of respondent error and/or misinterpretation. In addition, the absence of reported difficulty regarding the interpretation or appropriateness of the response scales used in the current study offers support for the use of these response options in the development of new population specific QoL scales.

It has been previously recognised that it may be easier to identify difficulties with questions than it is to fix these difficulties. Cognitive pre-testing can only serve to identify, not resolve or amend problem items as this is the role of the researcher (Willis, 2005). In this respect, when problems with items were identified, returning to the procedures undertaken for phase one of research and undertaking iterative peer de-briefing sessions undoubtedly facilitated the item modification process. Although of equal importance to this process was the clarity with which respondents reported their recommendations which resulted in high quality qualitative data. Such clarity of data a) increased the likelihood that recommendations were interpreted appropriately by the research team, b) facilitated the production of reports that documented the item modifications recommended by respondents and c) supported the clarity of data communicated to the research team for an amendment consensus (Spall & Stephen, 1998; Willis, 2005). Incorporating structured pre-testing protocols that identified potential problems with items prior to pre-testing combined with both think-aloud and probe questioning techniques facilitated the clarity of data reported. For example, the probing protocols ensured that an adequate amount of attention was given to assessing the performance of test items and that potential item failures where not overlooked. Additionally, in cases where respondents struggled to think-aloud or the frequency with which they undertook this task reduced, introducing additional probes helped to generate ongoing feedback from respondents.

3.10.1 Limitations of the Research

It is worth noting some of the characteristics of the participant sample for the current study. It is recognised that in isolation, the fact that the majority of respondents were female may well pose a limitation in terms of the scope for the utility of the scale under development. However, this female gender bias is representative of the populations that are referred into UK schemes and has been a finding of other studies that have utilised exercise referral participants for the purposes of research (e.g., Crone et al., 2005; Harrison et al., 2005; Lee, Griffin, & Simmons, 2009). It was also considered that the larger, demographically broader sample sizes that were included in the phases of scale reliability and validity procedures that followed would provide greater insights into the strengths and limitations of the utility of the scale in particular. Future research that is planned to further assess the performance of the new scale (see Chapter VII) will also provide opportunities to address gender equity.

3.11 Conclusions

A two phase approach to the early phases of scale construction at the item level proved effective in meeting the research aims. Specifically, phase one that made use of previous publications (e.g., Brace, 2004; Foddy, 1993; Hague, 1993; Oppenheim, 1992; Streiner & Norman, 2008) to inform effective item construction, questionnaire formatting and general presentation played a key role in generating a robust test pool of 50 items which were also subject to an iterative phase of peer de-briefing (Spall & Stephen, 1998). However, phase two of research (cognitive pre-testing) revealed problems with eight of the 50 items which, given the methods employed to ensure the construction of robust items during phase one, highlights the importance of including this second phase of item construction which is sometimes omitted by other researchers. The implications for the phases of research to follow are that cognitive pre-testing limited the likelihood of error due to question misinterpretation (De Leeuw, 2001; Drennan, 2003).

The response scales used for the current test measure were developed from previous research that specifically addressed response scale option wording within the context

of health scale development (Skevington & Tucker, 1999). Respondents identified no problems with the interpretation of the response scales used and found them well matched to items and in this respect the current study offers support for the use of these response options for similar scales under construction.

The methods employed in the current study and the level of detail reported have evidenced good face validity and may also contribute to the clarification of the methods available to scale construction researchers who are met with uncertainty regarding the procedures to be undertaken for the purposes of cognitive pre-testing (Presser, 2004).

Following pre-testing, a complete amended version of the test questionnaire was prepared (Appendix 14). This test measure was then subject to psychometric procedures of item analyses, reduction and initial reliability which are reported in the following chapter (Chapter IV).

CHAPTER IV Phase Three: Item Analyses, Reduction & Initial Reliability

4.1 Introduction: Item Analyses, Reduction and Reliability

Following pre-testing the next phase of research was designed to statistically test the performance of the 50-item test pool so as to identify those items that performed the best and would be retained for the final psychometric phase of research (Chapter V). This procedure is commonly referred to as item analysis and is typically conducted early on in the analytical/psychometric procedures during scale development. Although the methods used to conduct item analysis vary, the primary function of the methods used in the current phase of research was to investigate the reliability of the measure and to reduce a large number of possible items that have been generated to a more manageable number (Rust & Golombock, 2009). Although there are common features to most item analyses there are "as many recipes for item analysis as there are for chilli" (Murphy & Davidshofer, 2001, p.202). However, in terms of theoretical groundings, item analytical procedures are often rooted in Classical Test Theory (e.g., Lord & Novick, 1968; Novick, 1966), Item Response Theory (e.g., Embretson & Reise, 2000; Lord, 1980) and Rasch Theory (e.g., Andrich, 1988; Fischer & Molenaar, 1995). These theoretical groundings informed decisions regarding the most appropriate methods for the current study and as such a critique of these approaches to item analysis is presented below.

Classical Test Theory (CTT) has been the foundation for measurement theory for over 80 years and has allowed for the development of psychometrically sound scales (Kline, 2005). However, some of the underlying principles of CTT have their limitations. CTT assumes that the standard measurement of error is consistent across the entire population. Additionally, regardless of the total test score (e.g., high, medium or low perceived QoL), the standard error for each score is the same and as tests become longer, they become increasingly more reliable (Kline, 2005; Streiner & Norman, 2008). Although as the number of test-items increases, so does the required sample size to assess the performance of each item and larger numbers of respondents make the statistics generated by that sample more representative of the population than would a smaller sample. Similarly, larger numbers of items and the statistics generated by them (e.g., mean scores) are deemed to be more stable (Kline, 2005). In general, CTT evaluates at least three properties of scale items; the number of people choosing each response option, the distribution of responses, and the

relationship of item responses to other items in the measure (Murphy & Davidshofer, 2001).

Item Response Theory (IRT) also referred to as Latent Trait Theory is a model-based measurement in which trait level estimates depend on both persons' responses and on the properties of items that were administered. Structural equation modelling (SEM) is often the analytical procedure used to mathematically explore the relationships between these variables termed 'latent variables' that are not directly observed (or those that are inferred). IRT assumes unidimensionality (i.e., the existence of a single trait or construct underlying a set of measures; Hattie, 1985) and as a consequence Streiner and Norman (2008) suggest that IRT cannot be used to construct indices where the items are causal rather than effect indicators. Steiner and Norman (2008) have suggested that mobility, dressing feeding and leisure opportunities are examples of causal indicators of QoL because such indicators are said to impact upon perceived QoL rather than develop as an effect of a given level of perceived QoL (causal and effect indicators are discussed in more detail in Chapter III). Of the 50 items that comprise the current test-measure pool, 43 were identified as being clear causal indicators of life-quality. IRT has also been identified as problematic for exploring underlying constructs that are multi-faceted and complex such as those that are typical of those in the health field (Streiner & Norman, 2008). For example, in a study that aimed to improve items to assess physical function (Bruce, Fries, Ambrosini, Lingala, Gandek, Rose et al., 2009), the authors asserted that a unidimensional model may not be a valid representation of hand disability. This echoes some of the considerations during item analyses (Chapter IV) regarding too higher specificity whereby items contained within a scale are limited in the breadth of traits or lifestyle factors assessed and therefore, are unable to capture the complete range of factors that contribute to that which it purports to measure (Cattell, 1973).

Rasch Theory - also referred to as the Rasch Model or Rasch Analysis is often considered as a type of IRT. Scientific investigation and psychological theory development typically involves an iterative process of the modification of parameters that are accepted or rejected based on how well they fit the data. In contrast, when the Rasch Theory is employed, the objective is to obtain data which fit the theoretical model (Andrich, 2004). In this way, Rasch Theory assumes an underlying construct rather than an exploratory approach. Adopting a Rasch approach to psychometrics

has received attention in the literature (e.g., Heesch, Masse & Dunn, 2006; Nijsten, Unaeze & Stern, 2006; Ramp, Khan, Misajon & Pallant, 2009). Indeed, the careful consideration of which item-analytical approach to adopt for psychometrics and scale development has been the source of much debate within scientific publications (e.g., Hambleton & Jones, 1993; Prieto, Alonso & Lamarca, 2003).

Specifically, Hambleton and Jones (1993) described standard test development techniques of CTT for item analysis as: a) determining sample-specific item parameters by employing simple mathematical techniques and moderate sample sizes, and b) deleting items based on statistical criteria. The authors described that standard item analytic techniques involve an assessment of item difficulty and discrimination indices. The authors suggested that an important concern for test developers who apply the CTT approach is to ensure that the test sample used to validate the measure is as representative as possible. This is not surprising perhaps if it is considered that item statistics derived in CTT, such as item difficulty and discrimination are dependent on the sample of respondents selected to answer the items. However, the notion of representativeness is not exclusive to psychometrics and arguably should be of primary concern to all scientific investigation. This is particularly relevant to the health sciences where advances in theory and measurement are typically aimed at improving individual and population health profiles.

Hambleton and Jones (1993) identified that despite inherent difficulties in obtaining a representative sample, an advantage of the CTT approach to item analysis is that item statistics can be accurately calibrated on examinee samples of a modest size. The authors suggest that between 200 and 500 participants are required for a CTT item analytical approach while, in general 500 or more may be required for IRT. It was also suggested that classical item analysis procedures have the potential to provide the test developer with invaluable information regarding test item quality regardless of which measurement model that is applied in the later stages of scale development.

While Hambleton and Jones (1993) provided some useful non-empirical discussion regarding the CTT and IRT approaches to item analysis, Prieto et al. (2003) compared CTT principles and the Rasch Theory approach to reducing items included in The Nottingham Health Profile (NHP; Hunt, McKenna, McEwen, Williams, &

Papp, 1981) quality of life questionnaire. CTT resulted in a reduction to 20 items with four dimensions whereas Rasch Analysis resulted in a reduction to 22 items and two dimensions. Despite differences in content, total scores for both reduced item measures showed high degrees of association (0.78-0.95). Importantly, Prieto and colleagues (2003) asserted that the results of their study do not provide evidence to prefer one of the reduced item measures over the other but that the four dimensions of the CTT reduced measure did not fit the Rasch model specification. This study highlights some of the statistical difficulties inherent in scale development. Where one method may result in certain items being retained, another may lead to their exclusion. It is this inconsistency in outcome perhaps that is responsible for item-analytical methodological debate and again highlights why these considerations were so crucial to the current study design.

Hambleton and Jones (1993) and Prieto et al. (2003) provide some insights into the strengths, limitations and rationale for approaches to item reduction. However, other authors have reported their findings of scale development with a clear description of the item reduction methodology implemented but with little or no justification for the chosen approach (e.g., Badia, Arribas, Ormaetxe, Peinado & Terreros, 2009; Bruce et al., 2009). It is reasonable that word limitations of scientific journals may contribute to the lack of detail reported in this respect. However, this may further contribute to some of the uncertainty that surrounds the debate between CTT, IRT and Rasch analytical approaches within the literature and for test developers who wish to devise a sound psychometric methodology. It appears that the strengths of IRT including Rasch Theory approaches lie in their application to model validation and item calibration. The opportunity to contribute to model-building using an IRT approach has been deemed as one of the important differences between this and classical methods (Wilson, Allen, & Li, 2006). However, it has been recognized previously (Hambleton & Jones, 1993) that an adequate fit of model-to data is essential for successful item analysis as items may appear to perform poorly as a consequence of an ill-fitted model. Perhaps one of the most useful functions of an IRT approach is that it allows the test developer to determine the contribution of each test item to the test information function independently of other items in the test. Although, one of the largest logistical limitations is the greater sample sizes required as compared to CTT.

The challenge in scale development and psychometrics therefore seems to be in mastering the balance between statistical outcomes alongside considering such issues as content validity, the suitability of any a-priori model that is used, population appropriateness and minimal respondent burden while creating a measure that is easy to use, relatively short in length and is clinically meaningful. In practice, research-focused health practitioners often have to consider what is statistically significant while maintaining clinical significance for specific patient populations as the two do not always coincide. In this respect, the role of content validity monitoring alongside statistical considerations, regardless of the approach utilized becomes an integral part of the research process.

Based upon the review of psychometric approaches presented here, for the current study, the exploratory focus of item reduction, the majority of items being causal indicators of life-quality (n = 43) and the absence of a theoretical a-priori model (save for the identification of a conceptual framework that was used to guide item redundancy) indicated that a CTT approach was the most suitable in meeting the study aims.

4.2 Aims

The immediate aim of item analyses was to reduce the item pool, rejecting items that failed to reach predetermined criteria. The longer-term aim of item analyses was to contribute to additional psychometric procedures (e.g., exploratory factor analysis; EFA) that when combined will statistically identify items suitable for the construction of a relatively short yet meaningful validated measure of exercise-related life-quality.

4.2.1 Objectives

Item analytical procedures rooted in a CTT approach have the capacity to meet a number of research objectives. For the current study, these were: a) to ensure that individual items within the measure are acceptable to respondents in that they are able to understand and respond to each item, b) to assess the score distributions for individual items within the measure, c) to assess the reliability of items within the

measure in terms of their reproducibility and internal consistency and d) to analyse the internal consistency of the shortened scale.

4.3 Methods

Ethical approval was obtained from Nottingham Trent University and the UK National Research Ethics Service (NRES; reference number 06/Q2404/36; date of approval 25/4/2006). All participants were provided with an NRES template information sheet describing the study (Appendix 15) along with a shortened version for quick reference (Appendix 16). Scheme personnel received a cover letter explaining the purpose of the research and thanking them for their participation (Appendix 17) and an administration protocol to assist them with appropriate recruitment (Appendix 18). All participants provided full written consent to their participation, again using the NRES consent form template (Appendix 3).

Throughout this phase of research there were a number of methodological considerations regarding the participants involved and how they were recruited, the materials required, most appropriate procedures and analyses. Each of these considerations is described below.

4.3.1 Participants

The data generation phase was conducted over a period of three months. Respondents were recruited from 16 referral schemes across England. The referral schemes that participated in the research programme were invited to do so via telephone. Of the 17 schemes that were approached to act as test-sites to generate the respondent data, only one declined due to resource limitations. Wales were not included in the study because at the time of data generation, Welsh schemes were undertaking a national evaluation of existing schemes in the area. Therefore, in the interests of respondent and scheme burden, schemes in this area were omitted from the study. Of the two schemes that were invited from the Scotland area, one had been involved in a large research study in the past and as a consequence had started to develop evaluation protocols, and the other reported that the scheme was in its infancy and would not be suitable to be included in the study. Northern Ireland were

not included in the study as the primary investigator did not have access to the details of schemes located here at the time the data were collected. Table 15 indicates the numbers of participants recruited at each site included in the study.

Table 15. Number of respondents recruited from each test site

Test Site	Numbers Recruited
Gedling (Nottingham)	5
Huntingdonshire	23
Peterborough	18
Mansfield (Nottinghamshire)	20
London SW1	14
Worcester	1
Newham	5
Sheffield	6
Plymouth	41
Bristol	25
Nottingham City	69
East Cambridgeshire	3
Great Yarmouth	6
Doncaster	10
Solihull	25
Enfield	7
Total	**278**

Exercise referral instructors from each recruitment site approached individuals who were either at the start, in the middle or who had been exercising for up to 12 months post-referral to ask if they would volunteer to complete the test questionnaire. In a similar item analytical study, 60 participants were recruited to assess the performance of a 44-item test pool (Trigg, Jones, & Skevington, 2007) which is considered sufficient for CTT procedures. However, given that a larger number of respondents was required for the subsequent EFA (~250; Cattell, 1978) data from 278 respondents was available and used for item analyses. It is recognised that adopting a self-selection approach to data collection has its limitations in terms of ensuring adequate representation of the sample. However, it has also been argued that the importance of potential bias generated from this method to participant selection is unclear; especially when it generally cannot be avoided (Keating, 1989). This was true of the approach to participant recruitment adopted for the current study. It was considered more detrimental to the experience of those referred for

exercise and too burdensome upon scheme resources to introduce the completion of the test measure as a compulsory part of scheme protocol.

For the purposes of test-retest analysis, 25 respondents completed the questionnaire twice, seven days apart. These 25 respondents were recruited by means of those who volunteered to complete the questionnaire twice rather than a single occasion. Participants attending 16 exercise referral schemes across England were approached and asked if they would be prepared to complete the test measure twice until 25 re-test responses were collected. This would ensure adequate power for the test-retest reliability analysis (Field, 2009). There has been some debate within the literature as to what length of time between the initial and follow-up test is the most appropriate for test-retest reliability. A very short lapse of time runs the risk of participants recalling their previous responses and therefore falsely increasing the correlation (Kline, 2000). Additionally, Marx (2004) has asserted that given test-retest reliability is defined as the consistency of scores obtained by the same individuals when re-examined with the same test on a different occasion, it must be measured in individuals who do not change between measurements. This can be problematic for scale development research targeting ill-health populations where symptoms may fluctuate or where participants are subject to an intervention designed to impact upon that which is attempting to be measured consistently. This was a particular area of consideration for the current study. It was considered that given the above factors, seven days was the most appropriate time lapse. A previous study that specifically sought to assess the test-retest reliability of the International Classification of Functioning, Disability and Health measure (ICF; WHO, 2001) utilised the same time gap. The specific procedures undertaken for the analysis of the test-retest data are documented in section 4.3.4.

The complete cohort of participants ranged in age from 19 to 87 years (M = 61, SD ± 12.81). A total of 20 respondents ages were missing in the returned data. A greater number of females (n = 148) than males (n = 128) were included in the analysis. Gender information for two respondents was missing. Respondents reported their ethnicity as White (n = 233), Mixed (n = 1), Indian (n = 2), Pakistani (n = 2), Other Asian (n = 4), Black Caribbean (n = 11), Black African (n = 2), Other Black (n = 1), Other Ethnic Group (n = 3). The ethnicity of 19 respondents was unknown.

Of the 278 respondents included in the analyses, the health reasons for referral were known for 215 respondents. The majority of respondents were referred for a single health condition (n = 147) while others were referred with two conditions (n = 55), three conditions (n = 10) and four conditions (n = 3).

The health reasons for which participants were referred for exercise were also typical of exercise referral populations. This is of particular relevance given that CTT principles advocate the use of a respondent sample that has similar relevant characteristics to those for whom the measure is intended (Rust & Golombok, 2009). Of the 147 respondents who were referred for a single health condition, the most frequently reported condition was coronary heart disease (CHD; n = 34; 12%) followed by weight loss (n = 27; 10%), myocardial infarction (MI; n = 21; 8%), hypertension (n = 18; 7%), improved fitness (n = 18; 7%), and diabetes (n = 13; 5%). Of the 55 respondents who were referred with two health conditions, the most frequently reported second condition was weight loss (n = 10; 4%), followed by hypertension (n = 7; 3%). Of the 10 respondents referred with three health conditions, the most frequently reported third condition was weight loss (n = 2; 0.7%) with an equal number reporting angina as a third health condition for which they were referred (n = 2; 0.7%). Of the three respondents referred with four health conditions, the most frequently reported fourth conditions were diabetes (n =1; 0.4%), hypertension (n =1; 0.4%) and musculoskeletal reasons (n =1; 0.4%).

The majority of respondents (54%) identified that they were at the midway stage of their referral programme (around six weeks). While 26% identified that they had completed their initial referral phase of exercise (12 weeks and beyond) and the least amount of respondents (20%) identified that they had just started their referral programme. Therefore, the majority of respondents (80%) had been exposed to at least six weeks of exercise and physical activity if not more.

Of the 25 respondents included in the test-retest analysis, participants ranged in age from 36 to 73 years (M = 60, SD ± 11.94). Only the age for a single respondent included in the test-retest analysis was missing. Almost an even number of females (n = 12) to males (n = 13) were included in the analysis. Respondents reported their ethnicity as White (n = 22), Pakistani (n = 1), Other Aisian (n = 1) and Unknown (n = 1).

The health reasons for a referral into exercise were known for all 25 test-retest respondents. Seventeen respondents were referred for a single health condition, while six of these individuals were also referred for two health conditions and a further two of the total cohort (N = 25) were referred for three health conditions. Of the 17 respondents referred for a single condition, the most frequently reported was CHD (n = 5; 20 %). The second most frequently reported condition was to improve fitness (n = 4; 16 %) and the third most frequent was weight loss (n = 3; 12 %).

Of the six respondents who were referred with two health conditions, diabetes was reported by two individuals (8 %) while hypertension, improvements in fitness, depression and sleep apnoea each accounted for a single report respectively (4 %).

Similarly to the complete participant cohort (N = 278) the majority of test-retest respondents reported that they were at the mid-stage of their referral (56%). Thirty-six percent reported that they had completed their initial referral phase of 12 weeks and eight percent of respondents accounted for those at the start of their referral.

4.3.2 Materials

A total pool of 50 test items were included in the analyses. Through focus group interviews with individuals who had completed at least 12 weeks of a referral into exercise (Chapter II) a conceptual framework for exercise-related QoL was developed which contained four conceptual domains of QoL and were representative of the qualitative data generated from the focus group analysis. Each of the 50 questions were grouped into these respective constructs or domains namely: leisure facilities and lifestyle physical activity (12 items), confidence and motivation (12 items), discomfort and fitness (4 items) and mental and physical well-being (22 items). The labels for each of the four conceptual domains were decided upon by considering what aspect of exercise-related life-quality the item had been constructed to represent (i.e., what aspect of the qualitative data were the item intended to target?). Following this, labels were drafted by the primary investigator that intended to capture a description of the items allocated to each construct. These draft labels were subject to peer de-briefing (Spall & Stephen, 1998) and a consensus was reached for labels that were considered to best represent the item group (construct). Figure 12 illustrates the conceptual framework and the number of items that represent each domain within the framework. Tables 16, 17, 18 and 19 indicate the

wording of each item and respective domain. The domains leisure facilities and lifestyle PA were combined because only a single item: “to what extent can you afford leisure facilities” refers to leisure facilities directly and this would have resulted in a single item domain which for a QoL scale is not desirable. It was also considered that type of and time at which to exercise may be related to those offered by leisure facilities and therefore were questions that related to leisure facilities indirectly. However, generic lifestyle PA factors could also be attributed to facilitators or perceived limitations of type of or time to exercise and as a consequence provided a further rationale to combine the domains leisure facilities and lifestyle PA. During peer-debriefing - when reviewing the items contained within this combined domain, the question structure also presented good flow and logical order which has been recognised as strength of item grouping (Foddy, 1993).

These domains were not considered as paramount to the progression of the development of the tool, nor were they upheld as the theoretical criteria by which items would be retained or rejected. Rather, the proposed domains provided a conceptual framework that facilitated the content validity process that, alongside CTT analytical procedures ultimately informed which items would be retained (Streiner & Norman, 2008).

Figure 12. Number of items representing the four conceptual domains of exercise-related life-quality

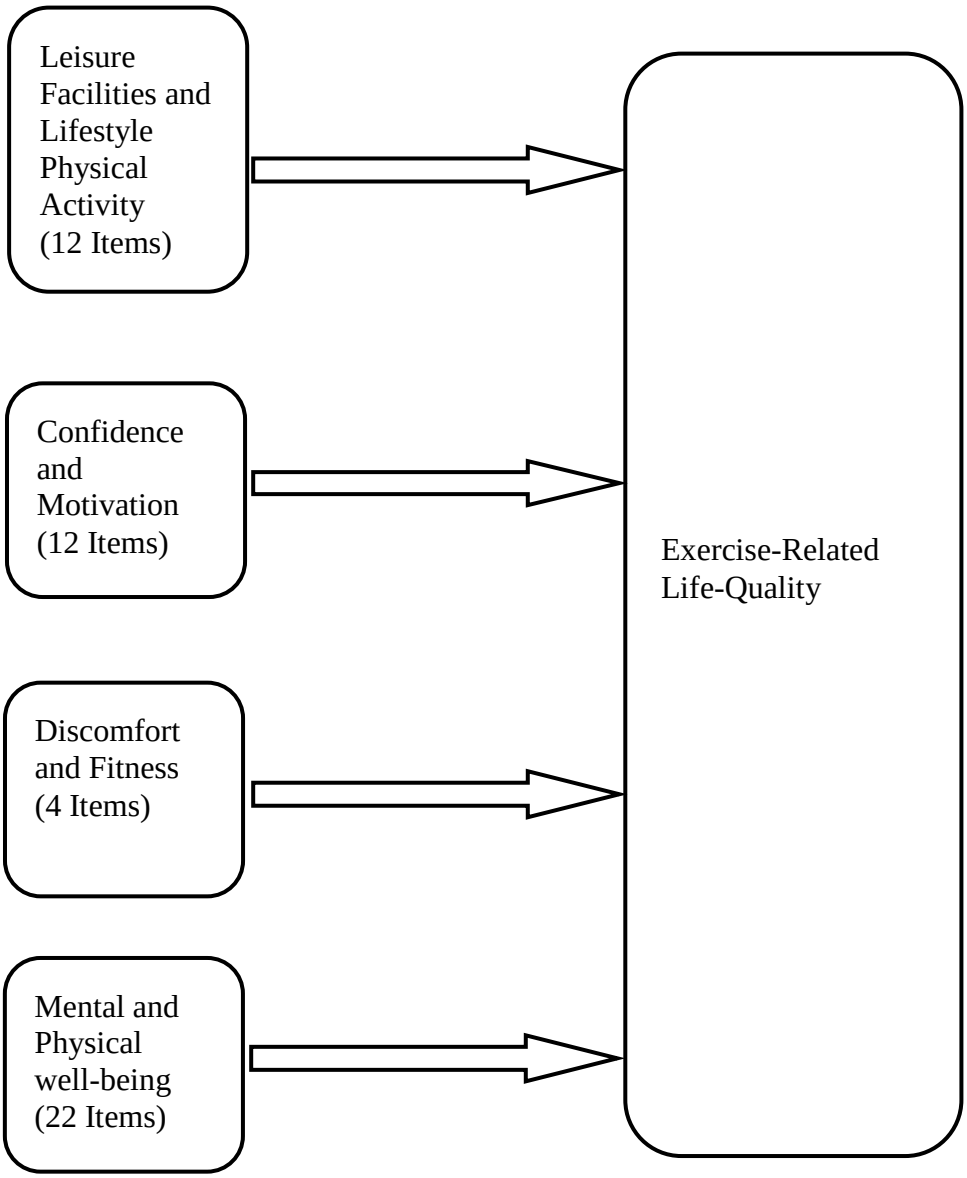


Table 16. Wording of items representing the construct of leisure facilities and lifestyle physical activity

Leisure Facilities and Lifestyle Physical Activity
Q1 To what extent can you afford leisure facilities
Q2) To what extent do you feel you have a choice regarding the type of exercise you undertake?
Q3) To what extent do you feel you have a choice regarding the time at which you exercise?
Q4) How much do you feel that lifestyle factors (e.g., transport, time, childcare, poor health & the weather) affect your ability to be physically active?
Q10) How much do you currently enjoy physical activity?
Q11) How much is social contact with others a part of your current lifestyle?
Q12) How important is it for you to manage your weight?
Q13) How much do you feel that you incorporate physical activity into your daily lifestyle?
Q14) How well do you feel you maintain eating habits that are beneficial to your health and any illness you may have?
Q15) In terms of exercise how competitive are you?
Q16) In terms of exercise how determined are you?
Q50) How would you rate your current knowledge of the benefits of physical activity and exercise for health?

Table 17. Wording of items representing the construct of confidence and motivation

Confidence and Motivation
Q17) How confident are you using gym equipment?
Q18) How confident are you that you can exercise in a leisure centre with minimum support?
Q19) How confident are you in your ability to participate in regular physical activity and exercise?
Q20) How confident are you in your ability to undertake daily tasks that do not necessarily involve physical activity?
Q21) In general, how confident are you around other people?
Q22) How confident are you in your own physical ability?
Q23) How confident are you to join in activities with family and friends?
Q24) How important is it to you right now to maintain a physically active lifestyle?
Q25) How confident are you, right now that you can maintain a physically active lifestyle?
Q26) How motivated are you to maintain a physically active lifestyle?
Q27) How likely are you to maintain a physically active lifestyle?
Q43) How would you rate your level of motivation to be physically active?

Table 18. Wording of items representing the construct of discomfort and fitness

Discomfort and Fitness
Q9) How much does any injury you may have prevent you from being physically active?
Q29) How often do you experience physical pain?
Q30) How often do you feel you have to manage any symptoms of pain?
Q31) How often do you feel you have to manage the symptoms of any illness you have?

Table 19. Wording of items representing the construct of mental and physical well-being

Mental and Physical Well-being
Q5) How much do you compare yourself to others while exercising?
Q6) To what extent do you worry about exercising in an environment with others?
Q7) To what extent do you worry about participating in structured exercise?
Q8) To what extent do you worry about the prospect of exercising in a fitness suite/gym?
Q28) In general, how much opportunity do you have to make time for yourself?
Q32) How often do you feel you have achieved something in terms of your physical fitness?
Q33) How would you rate your energy levels?
Q34) How would you rate your outlook on life
Q35) How would you rate your level of self-esteem?
Q36) How would you rate your physical health?
Q37) How would you rate your mental health?
Q38) How would you rate your sense of well-being?
Q39) How would you rate the quality of support you receive from leisure staff to be physically active?
Q40) How would you rate the quality of support you receive from family and friends to be physically active?
Q41) How would you rate the quality of your sleep?
Q42) How would you rate your memory function?
Q44) How would you rate your ability to manage stress?
Q45) How would you rate your ability to manage your mood?
Q46) How would you rate your ability to manage your weight?
Q47) How would you rate your current life expectancy?
Q48) How would you rate your ability to undertake everyday tasks that require some level of physical activity?
Q49) How would you rate your overall quality of life?

4.3.3 Procedure

Item analyses followed the procedures undertaken for the development of two previously validated scales: 1) a generic QoL scale, the WHOQOL-BREF (Skevington et al. 2004) and 2) the BASQID, an assessment of subjective QoL in dementia (Trigg, et al., 2007). Participants were provided with an information sheet following the format required by the National Research Ethics Service (NRES). Using the same format, written informed consent was also obtained from all participants.

Sixteen UK referral schemes returned completed questionnaires to the primary investigator via post reply. The primary aim of the current research was to assess the performance of the new scale at item level. Therefore, at this stage it was important to include those items that were shown to have missing responses as this was part of the statistical assessment criteria. In cases where data were missing in this respect, item responses were entered into SPSS (v17.0) as a value of 999 so that the frequency of missing data at item level could be easily identified. Sijtsma and van der Ark (2003) recognised that respondents may simply unintentionally miss a whole page of items. This is perhaps more likely for lengthy questionnaires such as the current test measure which contained a total of 88 items. In this instance, questionnaires are often described as incomplete rather than having occasional missing responses and previous studies have opted to remove these from the analyses as it is likely that the lack of completion is due to human error rather than item difficulty (e.g., Lejuez, Simmons, Aklin, Daughters & Dvir, 2004). Therefore, for the purposes of the current study, where respondents had overlooked a section of the complete 88 test items, these questionnaires were omitted from the data pool (n = 6). The confidentiality of respondents was ensured by translating participants' names into a number code (e.g., 01, 02, 03 etc.) Once the data were entered, questionnaires were stored in a locked cabinet to which only the primary investigator had access.

4.3.4 Analyses

In the main, data were managed and analyses conducted using SPSS (v17.0). For the purposes of test-retest analysis, weighted Kappa (*k)* scores were calculated using Microsoft Excel with add-on software Analyse-it. Data entry checks were made by

sampling random sets of questionnaires and cross referencing the responses given with the data that had been entered. All reverse-scored items for the initial larger 50-item pool (n = 9) and the subsequent, reduced 29-item pool (n = 6) were accounted for in the analysis so that all items in both pool data sets were computed in the same way.

The complete item analysis was an iterative process that included cyclic phases of peer de-briefing (Spall & Stephen, 1998) with the research team and the use of 'expert opinion' (Streiner & Norman, 2008) via consultations with two other members of the research team who were experienced in psychometrics and psychological tool development.

Frequency analyses were performed to assess response distributions at item level. Facility index (mean scores) of items were reviewed and those items that failed to fall within a mean range of 2-4 failed on this criteria. The range of responses (1-5) were also considered to assess how many of the response options were used by respondents. Those items that did not utilise the full range of response options failed this criteria. In line with the procedures developed by The WHOQOL Group (1998) and undertaken by Skevington et al. (2004), problematic items were also identified as those where the response distribution was skewed such that "any items with two or more adjacent scale points showing <10% of the responses on aggregate were highlighted as having frequency problems" (WHOQOL Group, 1998 p.1572). A question where not all of the response choices are utilised by respondents is undesirable, as this will reduce the sensitivity of the question in that it will be less able to discriminate between individuals with small differences in levels of the outcome being assessed.

Internal consistency was assessed using Cronbach's α and the contribution of each item (N = 50) to the total α. The corrected item-total correlations were also calculated. Kline (1987) suggested that as a general rule, an item should correlate with other items and the total score above 0.2. However, Skevington et al. (2004) identified poor item-total correlations in their development of the WHOQOL-BREF by considering those that were less than 0.3. Therefore, the item-total correlation level that was used to identify poor item-total correlations for the current study was <0.3. Finally, weighted kappa *(k)* was used to assess the test-retest reliability of each

item across one week. The strength of agreement between the responses using kappa was defined as poor (<0.2), fair (0.21-0.4), moderate (0.41-0.6), good (0.61-0.8), and very good (0.81-1.0) (Landis & Koch, 1977).

A priori criteria were set to guide item retention decisions. Items that failed all five of the item analytical criteria (i.e., facility index score, range consideration, corrected item-total correlation, adjacent response score and weighted kappa) were rejected without content validity consideration because such items had performed so poorly. All failed items were considered individually in terms of how many criteria had failed to have been met. The higher the number of failed criteria, the greater consideration was given to rejecting the item. In cases whereby a number of items represented a related aspect of exercise-related life-quality and all had failed, the item or items that had performed least well were rejected in favour of those that had statistically performed better. Similarly, items that failed a single criterion were rejected if the aspect of exercise-related life-quality that the item was designed to target was represented adequately by other items that performed better statistically. The size of the item pool (N = 50) was relatively large and therefore, those items that failed marginally on a single criterion were only retained if the aspect of exercise-related life-quality or the domain to which the failed item belonged was under-represented by other items.

Content validation considerations were monitored throughout the item selection process (Hambleton & Jones, 1993; Streiner & Norman, 2008). Once a combination of statistical and content validity considerations had identified those items that were rejected, Coronbach's α and revised item-total correlations were re-calculated for the reduced item pool.

4.4 Results

Item analyses of the item pool highlighted problems with 30 of the 50 items. A summary of the main findings of item analysis are presented in Table 20. The frequency of missing data were lowest (zero) for items: Q1 regarding affordability, Q4 lifestyle factors, Q6 worry regarding exercising with others, Q11 social contact, Q21 confidence around others, Q23 confidence regarding family and friends, Q26 motivation to maintain a physically active lifestyle, Q28 time for self, Q33 energy levels, Q37 mental health, Q43 PA motivation level, Q48 ability to undertake

physically active tasks, Q49 overall QoL and Q50 knowledge regarding the benefits of PA. Data were most frequently missing for Q17 confidence regarding the use of gym equipment (n = 8), Q29 frequency of pain experience (n = 6) and Q30 frequency of pain management (n = 6).

Facility index scores revealed seven items that failed to meet the mean range (2-4) criteria. These items were: Q6 worry regarding exercising with others, Q7worry regarding structured exercise, Q8fitness suite/gym worry, Q24 importance to maintain a physically active lifestyle, Q37 mental health, Q39 support from leisure centre staff and Q50 knowledge regarding the benefits of PA.

An analysis of the distribution of responses revealed 20 items that had adjacent scale points showing < 10% of the response on aggregate. The adjacent responses column of Table 20 indicates those responses (1-5) that were not endorsed by respondents at such a level so as to contribute to >10% of the total responses selected by all respondents for this item. The scale values (1-5) that failed this criteria are documented in this column (Table 20). The majority of these items (n = 18) were at the lower end of the scale (i.e., response scores of 1-2) rather than at the higher end of the scale (i.e., 4-5). Only two items; Q6 worry regarding exercising in an environment with others and Q8 worry regarding exercising in a fitness/suite gym environment had adjacent scale points showing <10% of the response on aggregate for responses four and five.

All items save for Q7, worry regarding structured exercise; Q13, incorporating physical activity into daily lifestyle; Q21, confidence around others; Q24, importance to maintain a physically active lifestyle; Q25, confidence to maintain a physically active lifestyle and Q50, knowledge regarding the benefits of physical activity utilised the whole range of response options (1-5). All aforementioned items that did not utilise the complete range of the response scale ranged in responses from 2-5 with the exception of item Q7 that ranged in responses from 1-4.

Internal consistency of the complete test item pool was excellent with a Cronbach's coefficient alpha of 0.92 (N = 50) and alpha could not be improved by omitting any single item. Eight items failed to meet the item-total correlation criteria level of >0.3 (Skevington et al., 2004). These items were: Q1 affordability (-.13), Q2, exercise

choice (.18) Q3, exercise time (.15) Q5 comparison to others (.24), Q11 social contact (.30), Q12 weight management (-.07), Q15 competitiveness (.24) and Q39 support from leisure staff (.29).

The test-retest reliability of items assessed using weighted kappa as defined by Landis and Koch, (1977) indicated that 48 items fell into the very good (0.81-1.0) category. Two items, Q16 determination level and Q24 importance to maintain a physically active lifestyle fell into the good category (0.61-0.8). The highest reported *k* value was for Q41 sleep quality (0.96). While the lowest reported *k* values were for Q16 and Q24 (0.78) respectively.

To summarise, item analyses of the item pool highlighted problems with 30 of the 50 items. Twenty failed on a single criterion, nine failed on two criteria and a single item failed on three criteria. No items failed more than three of the five criteria. See Table 20 for a detailed breakdown of which items failed each criteria.

Based upon the recommendations of Hambleton and Jones, (1993) and Streiner and Norman, (2008), content validity was monitored alongside item analytical outcomes which helped to inform item retention decision making. Those items that had failed the least number of statistical criteria within a given category were given priority consideration for retention in instances where item removal would challenge the broad representativeness of the complete scale. As a consequence, nine of the items that failed statistically were retained, resulting in a total of 21 test items being rejected. What follows is a detailed rationale regarding the retention of these nine items that failed statistically.

Q1 regarding the affordability of leisure facilities was constructed to allow for the assessment of how much the cost implications of attending leisure facilities may impact upon attendance. If an individual is unable to access a health-enhancing service due to financial limitations, it follows that this could impact negatively upon perceived QoL. It is commonplace for income-related items to be included in a QoL measure (e.g., WHOQOL-BREF). However, the specific reference to the affordability of leisure facilities contained within Q1 was intended to address exercise-related life-quality in particular.

Q1 was re-worded from “how important to you is the affordability of leisure facilities in order to participate in a structured exercise programme?” to read “to what extent can you afford leisure facilities?” The original wording of item Q1 failed on contributing to <10% of responses on aggregate and also on item-total correlation (-.13). The negative item-total correlation for this question (-.13), suggests that the item could be considered as a reverse scored item. However, the re-wording of this item should clarify the rationale for scoring in a positive direction because it is reasonable to suppose that a greater ability to afford leisure services should be positively associated with perceptions in life-quality. This was the only item that referred to perceptions of the affordability of leisure services. The final 26-itemWHOQOL-BREF includes a more general question regarding personal finances. The question reads: “to what extent do you have enough money to meet your needs?” The 44-item test pool of the BASQID also asked “how satisfied are you with your financial situation?” Although this item was subsequently omitted from the final 14-item measure. With respect to the current scale, because Q1 is the only item targeting affordability and because item analysis is conducted early in the psychometric process a revised question was developed using expert opinion (Streiner & Norman, 2008). Given that the WHOQOL-BREF has undergone appropriate item-analytical procedures to investigate item performance, it is feasible that using similar wording for the current measure will support appropriate levels of reliability and validity. However, further psychometric procedures (i.e., EFA) may indicate the need to remove this item from the final measure and similar to the final 14-item BASQID, the current final measure may not include a question related to the perceived affordability of leisure services.

Q5 comparison to others was also retained because - of the four items designed to assess perceived exercise anxiety within different contexts (of which Q5 was one) this item failed on a single item analytical criteria (corrected item-total correlation .24). However, the remaining three items (Q6 worry regarding exercising with others, Q7 worry regarding participating in structured exercise and Q8 worry regarding exercising in a fitness suite/gym) each failed on two item analytical criteria. Q6 and Q8 both failed the facility index and adjacent response criteria while Q7 failed the facility index criteria and did not utilise the full range of response options (1-5) instead only covering the range (1-4). In the interests of ensuring that

exercise-related anxiety was represented in at least one context, the item that performed most well statistically was retained.

Q10 enjoyment of PA was retained because the literature suggests that behaviours that are perceived as being worthwhile are more likely to be attempted or adhered to (Skinner, 1938). In this case, if PA is adhered to, the literature also suggests that exercise has a positive impact upon perceived life-quality (e.g., Bowen et al., 2006; Stathi et al., 2002). Q10 was the only item representing perceived enjoyment of PA and failed on a single criteria (adjacent responses 1 and 2). Therefore, scientific understanding regarding the role of the perceived benefit and enjoyment of a given behaviour, the documented relationship between exercise and enhanced life-quality and content validity considerations overruled the statistical criteria on this occasion.

Q11 regarding social contact failed to meet the item-total correlation criteria of >0.3 (.30). However, the focus group data (Chapter II) indicated that the prospect of social contact was often the source of motivation for attending a leisure activity and participants found solidarity in the knowledge that others had been referred for exercise. For example, one focus group participant commented I've made a lot of friends" and another "I thought I was on my own and then you go to the doctors and the nurse refers you to something like this and you realise you aren't on your own." Q11 was the only item representing social contact and given that the qualitative data highlighted the importance of social contact in the maintenance of exercise referral attendance this item was retained to increase the test measures representativeness of factors related to exercise-related life-quality. Furthermore, had the item-total correlation criteria been set to the lesser value of >0.2 (Kline, 1987) this item would have been retained. The higher criteria value of >0.3 resulted in this item marginally failing.

The eating habits of respondents, more particularly in terms of habits that are congruent to good health were represented by a single item: Q14 "how well do you feel you maintain eating habits that are beneficial to your health and any illness you may have?" This item failed on the adjacent response criteria whereby respondents did not make sufficient use of response options 1 (not at all; valid percent 1.1) and 2 (not much; valid percent 8.7). In the interests of developing a scale that has the capacity to address the breadth of exercise-related life-quality and avoid a measure that is too narrow with bloated specifics (Cattell, 1973). The item was retained

because no other item addressed nutritional habits. However, as with all items retained on the grounds of content validity, careful attention will be given to this item during subsequent analytical procedures.

Q34 regarding perceived outlook on life failed to meet the adjacent response criteria whereby respondents did not make sufficient use of response options 1 (very poor; valid percent 1.1) and 2 (poor; valid percent 4.7). Q34 coupled with Q37 "how would you rate your mental health?" and Q38 "how would you rate your sense of well-being?" sought to address perceptions of mental health and mental well-being more directly. However, items Q35 regarding self-esteem, Q41 regarding sleep quality, Q42 regarding memory function Q44 regarding stress management and Q45 regarding mood management address mental health and well-being more indirectly all of which fulfilled the item analytical statistical criteria. Q34 failed on a single statistical criterion. Similarly, Q38 regarding perceived well-being failed to meet a single criterion: the adjacent response criteria 1 (very poor; valid percent .4) and 2 (poor; valid percent 6.9) while Q37 failed to meet the mean score range criteria (4.06) and also did not make adequate use of response options 1 (very poor; valid percent 1.4) and 2 (poor; valid percent 3.2). As a consequence, in the interests of retaining items that addressed mental health and well-being more directly, Q34 and Q38 were retained at this stage of scale development because they failed fewer criteria than Q37, within this item group. The removal of Q34, Q37 and Q38 would have resulted in the absence of the scale's ability to assess mental health using more direct language (i.e., outlook on life, mental health and well-being). Therefore, again, the rationale of the inclusion of Q34 and Q38 at this early stage of scale development was to promote the scale's potential for the breadth of assessment of exercise-related life quality

Q47 regarding perceived life-expectancy failed to meet the adjacent response criteria whereby respondents did not make sufficient use of response options 1 (very poor; valid percent 1.4) and 2 (poor; valid percent 4.3). For those diagnosed with chronic disease, life expectancy is an important consideration (e.g., Stewart, Cutler & Rosen, 2010). Q47 was the only item that was designed to assess perceptions of life-expectancy and for these reasons this item was retained at this early stage of scale development.

Perceptions of QoL was addressed directly by Q49 “how would you rate your overall quality of life?” This item failed to meet the adjacent response criteria response 1 very poor (1.4) and 2 poor (4.7). However, because this was the only item that specifically refers to overall life-quality, at this stage of scale development, it was considered that there was value in including the item. The final measures of both the WHOQOL-BREF and BASQID include a question specifically referring to QoL.

Prior to item analyses four conceptual item domains were used to inform item retention: leisure facilities and lifestyle physical activity (n =12), confidence and motivation (n = 12), discomfort and fitness (n = 4) and mental and physical well-being (n = 22). Following item analyses the items retained in each category were: leisure facilities and lifestyle physical activity (n =5), confidence and motivation (n = 4), discomfort and fitness (n = 4) and mental and physical well-being (n =16). Cronbach’s α for the revised item pool (n = 29) was good (0.89). Of the nine items included in the initial larger test pool (N = 50) that failed statistically but were retained on the grounds of content validity, four of these retained items failed again in the reduced pool (N = 29) failing to reach an item-total correlation value of >0.3. These items were: Q1 regarding affordability, Q5 regarding comparisons to others, Q11 regarding social contact and Q14 regarding eating habits. These items will continue to be retained at this stage in the scale development process for the same content validity reasons detailed in the results of the initial larger item-pool analysis. However, these items will be subject to careful consideration during the EFA that will follow.

The final set of 29 items, the domain to which they belong and revised item-total correlations can be reviewed in Table 21. Of the nine reverse scored items included in the test pool, six were retained. These were Q4 regarding lifestyle factors, Q5 regarding comparisons to others, Q9 regarding injury prevention, Q29 regarding the frequency of pain, Q30 regarding the frequency of pain management and Q31 regarding illness symptom management**.**

Table 20. Summary of item analysis of the test-pool (N = 50)

Item	Valid n index	Facility	Range	Corrected item total correlations	Adjacent responses	Valid percent	Weighted kappa (n = 25)
*Q1	278	3.91	1-5	**-.13**	**1.00**	**1.4**	0.91
					2.00	**5.0**	
Q2	274	3.84	1-5	**.18**	**1.00**	**1.1**	0.92
					2.00	**8.0**	
Q3	277	3.52	1-5	**.15**			0.83
Q4	278	3.00	1-5	.45			0.88
*Q5	275	2.35	1-5	**.24**			0.90
Q6	278	**1.67**	1-5	.41	**4.00**	**3.2**	0.92
					5.00	**1.8**	
Q7	277	**1.67**	**1-4**	.47			0.89
Q8	277	**1.76**	1-5	.44	**4.00**	**6.9**	0.92
					5.00	**.7**	
Q9	275	2.51	1-5	.46			0.88
*Q10	274	3.81	1-5	.51	**1.00**	**.4**	0.87
					2.00	**5.5**	
*Q11	278	3.49	1-5	**.30**			0.89
Q12	276	3.79	1-5	**-.07**	**1.00**	**5.1**	0.93
					2.00	**4.3**	
Q13	276	3.36	**2-5**	.38			0.92
*Q14	277	3.53	1-5	.35	**1.00**	**1.1**	0.89
					2.00	**8.7**	

Continued

Item	Valid n	Facility index	Range	Corrected item total correlations	Adjacent responses	Valid percent	Weighted kappa (n = 25)
Q15	276	2.67	1-5	**.24**			0.94
Q16	275	3.78	1-5	.32	**1.00** **2.00**	**1.1** **1.5**	0.78
Q17	270	3.50	1-5	.40			0.93
Q18	276	3.46	1-5	.40			0.89
Q19	277	3.62	1-5	.54	**1.00** **2.00**	**.7** **7.6**	0.89
Q20	277	3.86	1-5	.41	**1.00** **2.00**	**.4** **7.2**	0.88
Q21	278	3.88	**2-5**	.51			0.95
Q22	275	3.41	1-5	.65			0.85
Q23	278	3.75	1-5	.60			0.87
Q24	277	**4.29**	**2-5**	.41			0.78
Q25	277	3.79	**2-5**	.67			0.89
Q26	278	3.96	1-5	.49	**1.00** **2.00**	**.4** **2.9**	0.88
Q27	277	3.81	1-5	.51	**1.00**	**.4**	0.91

Continued

Item	Valid n	Facility index	Range	Corrected item total correlations	Adjacent responses	Valid percent	Weighted kappa (n = 25)
Q28	278	3.38	1-5	.33	**2.00**	**3.2**	0.90
Q29	272	2.77	1-5	.38			0.83
Q30	272	2.67	1-5	.44			0.88
Q31	275	2.97	1-5	.31			0.85
Q32	275	3.29	1-5	.43			0.82
Q33	278	3.39	1-5	.54			0.90
*Q34	277	3.98	1-5	.53	**1.00**	**1.1**	0.93
					2.00	**4.7**	
Q35	277	3.72	1-5	.59			0.95
Q36	276	3.31	1-5	.64			0.95
Q37	278	**4.06**	1-5	.56	**1.00**	**1.4**	0.91
					2.00	**3.2**	
*Q38	276	3.79	1-5	.64	**1.00**	**.4**	0.95
					2.00	**6.9**	
Q39	277	**4.43**	1-5	**.29**	**1.00**	**.4**	0.84
					2.00	**.4**	
Q40	277	3.99	1-5	.32	**1.00**	**1.8**	0.91
					2.00	**4.3**	
Q41	277	3.23	1-5	.35			0.96
Q42	276	3.53	1-5	.39			0.92

Continued

Item	Valid n	Facility index	Range	Corrected item total correlations	Adjacent responses	Valid percent	Weighted kappa (n = 25)
Q43	278	3.80	1-5	.48	**1.00** **2.00**	**.4** **6.1**	0.92
Q44	276	3.53	1-5	.50			0.91
Q45	274	3.60	1-5	.46			0.93
Q46	275	3.31	1-5	.46			0.87
*Q47	276	3.74	1-5	.54	**1.00** **2.00**	**1.4** **4.3**	0.91
Q48	278	3.58	1-5	.61			0.93
*Q49	278	3.90	1-5	.60	**1.00** **2.00**	**1.4** **4.7**	0.92
Q50	278	**4.17**	**2-5**	.49			0.87

Note. Figures in bold indicate items that failed the respective item analytical criteria. Only items that failed to meet the adjacent response criteria are reported in bold. *Denotes those items that failed statistically but were retained on the grounds of content validity.

Table 21. Retained items, domains and revised item-total correlations (N = 29)

Item	Domain	Revised Item-Total Correlation
Q1) How important to you is the affordability of leisure facilities in order to participate in a structured exercise programme?	Leisure and lifestyle PA	**-.17**
Q4) How much do you feel that lifestyle factors (e.g., transport, time, childcare, poor health & the weather) affect your ability to be physically active?	Leisure and lifestyle PA	.44
Q5) How much do you compare yourself to others while exercising?	Mental and physical health	**.22**
Q9) How much does any injury you may have prevent you from being physically active?	Discomfort and fitness	.47
Q10) How much do you currently enjoy physical activity?	Leisure and lifestyle PA	.45
Q11) How much is social contact with others a part of your current lifestyle?	Leisure and lifestyle PA	**.28**
Q14) How well do you feel you maintain eating habits that are beneficial to your health and any illness you may have?	Leisure and lifestyle PA	**.28**
Q17) How confident are you using gym equipment?	Confidence and Motivation	.32
Q18) How confident are you that you can exercise in a leisure centre with minimum support?	Confidence and Motivation	.36
Q22) How confident are you in your own physical ability?	Confidence and Motivation	.61
Q23) How confident are you to join in activities with family and friends?	Confidence and Motivation	.55
Q28) In general, how much opportunity do you have to make time for yourself?	Mental and physical health	.32
Q29) How often do you experience physical pain?	Discomfort and fitness	.40
Q30) How often do you feel you have to manage any symptoms of pain?	Discomfort and fitness	.48

Continued

Item	Domain	Revised Item-Total Correlation
Q31) How often do you feel you have to manage the symptoms of any illness you have?	Discomfort and fitness	.35
Q32) How often do you feel you have achieved something in terms of your physical fitness?	Mental and Physical Health	.37
Q33) How would you rate your energy levels?	Mental and physical health	.57
Q34) How would you rate your outlook on life?	Mental and physical health	.57
Q35) How would you rate your level of self-esteem?	Mental and physical health	.58
Q36) How would you rate your physical health?	Mental and physical health	.67
Q38) How would you rate your sense of well-being?	Mental and physical health	.67
Q41) How would you rate the quality of your sleep?	Mental and physical health	.38
Q42) How would you rate your memory function?	Mental and physical health	.37
Q44) How would you rate your ability to manage stress?	Mental and physical health	.51
Q45) How would you rate your ability to manage your mood?	Mental and physical health	.47
Q46) How would you rate your ability to manage your weight?	Mental and physical health	.45
Q47) How would you rate your current life expectancy?	Mental and physical health	.55
Q48) How would you rate your ability to undertake everyday tasks that require some level of physical activity?	Mental and physical health	.62
Q49) How would you rate your overall quality of life?	Mental and physical health	.65

4.5 Discussion

The initial item pool contained 50 items developed from four different conceptual domains of exercise-related life-quality (i.e., leisure facilities and lifestyle physical activity, confidence and motivation, discomfort and fitness and mental and physical well-being). Of the 50 items analysed using CTT, 30 items failed one or more item-analytical criteria. However, nine of these were retained on the grounds of content validity. It was recognised that 50 test items that required responses could potentially have posed a significant burden upon participants. Particularly for those who volunteered to complete the test measure on two separate occasions for the purposes of test-retest analysis. For this reason, the four conceptual domains of exercise life-quality that were identified pre-analysis that were used to guide the process of item retention and ensure an appropriate level of content validity, often had a single item that represented the areas that contributed to each respective domain. While including a greater amount of items may have assisted with selecting the most appropriate question, including more was not feasible for respondent and leisure centre staff resource burden. This limitation offers further support for the importance of cognitive pre-testing that preceded item analyses alongside content validity and expert opinion (Streiner & Norman, 2008) that also helped to establish the performance of question wording and interpretation during item analyses.

The current QoL scale under development has been designed to be sensitive to exercise referral. However, it is worth critiquing the similarities and differences between the current scale and the generic measure of QoL the WHOQOL-BREF in terms of item-level analyses because this will provide some insights into the specificity, suitability and initial performance of the scale under construction. The item analytical methods used in the current study were also based upon those employed by the WHOQOL-BREF and also the BASQID. However, it is more relevant to draw comparisons between the WHOQOL-BREF and the current scale rather than the BASQID at this stage as the BASQID was designed specifically for dementia. Therefore, the development of the BASQID is only relevant in terms of the CTT item analyses methods employed to its construction which were used to inform the current study.

Internal consistency (Cronbach's α) for the initial larger item pool (N = 50) (0.92) was consistently higher than sample values reported for the WHOQOL-BREF (Skevington, et al., 2004) with the highest score reported for the WHOQOL-BREF being for the dimension of physical health (0.82). However alpha is a function of item numbers – the more items the higher the value of alpha is likely to be. The current test measure had a greater number of test pool items than the WHOQOL-BREF. Skevington et al. (2004) reported that poor item-total correlations (<0.3) were only found for the dimension of negative feelings and only in one of the test centres used to generate the data for the development of the WHOQOL-BREF.

The test items for the WHOQOL-BREF were developed from the longer WHOQOL-100 within the context of four domains: physical, psychological, social and environment (Skevington, 1999). As a consequence, the rejection of items from the WHOQOL-BREF pool are not reported in the same manner as the current study because the latter two measures were developed as new measures of self-report life-quality. The resultant retained 29 items (58%) are comparable to the numbers included in the WHOQOL-BREF (26 items) final scale. The final number of items (N = 29) is a favourable number of items to move forward into the second phase of scale psychometrics – EFA. The number of items is not so small as to run the risk of reducing the item pool to such an extent that the measure would be limited in its ability to adequately assess exercise-related life-quality. EFA may help to further reduce the item pool and assist in developing a measure that is suitable for administration in exercise referral settings where time is limited and respondent burden should be minimised.

As a method, the procedures employed by CTT proved to be valuable in assessing the performance of items and reducing the item pool. Without these procedures, even with prior cognitive-pre testing the task of selecting items that best represent exercise-related life-quality and that are most relevant to the exercise referral population would have been undesirably based on tool developer intuition and item preference alone. The data set was large enough (278 respondents) to make confident decisions regarding statistical performance and the item analytical criteria set was based upon sound previous research. However, it is interesting to note that had the item-correlation criteria been set at a lower recommended level of >0.2 (Kline, 1987), four of the items included in the initial larger-pool (N = 50) that failed the corrected item-total correlation criteria would have fulfilled this criteria. Of these

four, three would have been retained in the item pool. This exemplifies the critical role of content validity considerations and the application of appropriate methods to item analyses that were based on prior research.

High internal consistency reliability is often considered a prerequisite of high validity (e.g., Guilford, 1956; Nunnally, 1978). However, items within a scale that are too highly correlated may indicate that a scale is too narrow in its focus. Kline (2000) noted that there are many tests whose items are paraphrases of each other which gives rise to high internal consistency but results in a scale that is highly specific. Cattell (1973) refers to the specificity of such tests as bloated specifics. The scale currently under construction is intended to assess related dimensions that also have the capacity to contribute something unique to the scale as a whole. Therefore, a Cronbach's α of 0.89 for the reduced item pool (N = 29) would suggest an ideal level of internal consistency. While a reliability of 0.7 is considered a minimum value (Kline, 2000) scale developers often strive to achieve internal consistency values close to or above 0.9. However, as Kline (2000) and Cattell (1973) have asserted, such values would indicate a measure that is highly specific containing variables with little breadth which would be undesirable for measures claiming to target aspects of perceived life-quality such as the scale currently under construction.

A greater number of females (n = 148) than males (n = 128) were included in the analysis. This appears to be typical of the gender weighting of populations that are referred into UK schemes, as exemplified by previous research (e.g., Crone et al., 2005; Harrison et al., 2005; Lee, Griffin, & Simmons, 2009). Similarly, participants in the current study ranged in age from 19 to 87 years (M = 61, SD ± 12.81). Previous studies have reported similar ages of participants recruited into exercise referral research. For example, Crone et al. (2005) reported a participant mean age of 63.2. The sample used in the current study supports the adult age range targeted by DH for PA improvement within primary care settings (DH, 2006) and in these respects the sample size is representative of those for whom the final measure is intended. This has been identified as a pre-requisite of the CTT approach (Kline, 2000) and should also be considered good practice as there would be little justification for a measure that is developed using very different participant population demographics than those for whom the measure is intended. That said,

there are examples of measures that have been psychometrically tested (at least at some phase of scale development) using student populations as these cohorts are often the most accessible to researchers and scale developers (e.g., Buck et al., 2005).

Proportionately more respondents reported their ethnicity as White (n = 233) with the second highest majority being reported as Black Caribbean (n = 11). Unless exercise-referral population studies are concerned with ethnicity specifically as a factor in exercise behaviour research, ethnicity data is rarely reported. Although, in a comprehensive RCT, Isaacs et al. (2007) reported greater numbers of white participants (n = 240; 75.7%) than Asian participants (n = 53; 16.7%) who had attended a leisure centre based exercise referral scheme however, no other ethnic categories were reported. Indeed, in a subsequent study that aimed to ascertain the factors associated with exercise referral uptake and participation, ethnicity was not reported (James et al., 2008). As a consequence, it is difficult to ascertain whether the sample used in the current study is representative of exercise-referral populations in terms of ethnicity. Although, given the geographic spread of schemes that recruited participants into the study, it is likely that respondents are representative of those for whom the final measure is intended.

Respondents referred with either one, two, three or four health conditions were referred for CHD or CHD risk factor related illness (e.g., weight loss/hypertension). This supports the health priorities for health improvement in the UK (DH, 2004c) and also perhaps highlights the rationale for the wealth of research that has investigated many different aspects of CHD such as: prevention (e.g., Thomsen, Davidsen, Ibsen, Jørgensen, Jensen, & Borch-Johnsen, 2001), economic burden in the UK (Liu, Maniadakis, Gray & Rayner, 2002), and the role of exercise in risk-factor reduction (e.g., Taylor et al., 2006). In practice, cardiac rehabilitation exercise classes are usually delivered separately to generic exercise referral classes and are often referred to as Phase IV cardiac rehabilitation classes. The general premise of Phase IV cardiac rehabilitation is that community dwelling heart patients are provided with the knowledge, skills and opportunities to choose health behaviors that are congruent to good heart health as a long-term, rest-of-life approach to rehabilitation. It is not unusual that the exit route for cardiac patients attending these specialized classes is to be offered opportunities for exercise that comprise those available to the exercise referral scheme in the locality. One of the challenges

therefore, is the continuity of evaluation from Phase IV exercise participation to exercise referral. Indeed, despite the patient referral pathway relationship between the two, Phase IV and exercise referral are often treated very separately. The health profiles of respondents used in the current study suggests that the final scale may have the capacity to collect exercise related life-quality data for exercise referral participants, cardiac patients attending Phase IV classes, and those who may have progressed from Phase IV classes into exercise referral. The potential for the final scale to be used in such a flexible manner should be considered a particular strength of the current study.

4.5.1 Limitations of the Research

The limitations of the current research lie largely within the demographics of the sample used. For example, ethnic groups other than White were less well represented and there were a greater number of females (n = 148) than males (n = 128) that were included in the analyses. As mentioned previously, a further limitation is inherent in the exclusion of exercise referral schemes based in Wales, Scotland and Northern Ireland because of logistical and feasibility reasons. Future work is planned to test the performance of the final scale (see Chapter VII) whereby there will be opportunity to address these sample demographic limitations.

4.6 Conclusions

Employing a CTT approach to item analyses and using previous CTT based item analytical research to guide the research process (i.e., Skevington et al., 2004; Trigg et al., 2007) proved an effective method of assessing the performance of the scale under development at item level and reducing the test item pool in preparation for further psychometric analyses. Coupled with the CTT approach the role of content validity played an integral part of the item retention process and helped to increase the probability of a final measure that is accurate in what it claims to measure (valid), is able to measure consistently (reliable) and has sufficient breadth of item content so as to avoid too greater level of specificity (e.g., Klein, 2000) or what has been termed bloated specifics (Cattell, 1973). In terms of health-related assessment, this approach

was also successful in assessing those items that performed well enough to represent such a multi-faceted and complex construct as exercise-related life-quality (Streiner & Norman, 2008).

The demographic profile of respondents was such that the sample used for item analyses is likely to be representative of the population for whom the final measure is intended and also reflects those of previous exercise referral research (e.g., Crone et al., 2005; Isaacs et al., 2007). Furthermore, the health reasons for which respondents were referred for exercise may result in the capacity for the final measure to be used across both Phase IV cardiac rehabilitation and exercise referral programmes.

Following item analyses the next step was to assess how the questions should be grouped together and also the reliability and validity of the items at scale level. The following chapter describes how the factor structure of the test measure was established including the internal consistency and test-retest (reliability), construct validity and acceptability of the resultant final scale.

CHAPTER V: Phase Four: The Factor Structure, Reliability, Validity and Acceptability of the Final Scale

5.1 Introduction: Factor Analysis

The previous phase of research described the processes undertaken to reduce the item pool so that only statistically rigorous items and those retained on the grounds of content validity were included in the current study. Save for the focus group, the previous phases of research have very much focussed on an item level of data construction and analyses. The next phase of research addresses how to combine these individual items into a scale which will inform how to express the final score in the most meaningful way (Streiner & Norman, 2008). Therefore the focus of the current research was to reduce the item pool further (N = 29) and to establish the factor structure of these remaining items. In order to meet this aim, factor analysis (FA) combined with an assessment of Scree plots and on occasion parallel analysis is often employed although there are different approaches to FA, primarily depending upon whether the focus of research is confirmatory or exploratory.

Establishing which approach to FA would be most suitable, how many factors to extract, rotation type, and additional analyses that could be used to help support the interpretation of the factor structure were critical considerations during the planning phases of the current research. Therefore, some relevant background and a critique of these considerations with respect to the aims of the current research is presented below. Additionally, a critique of the methods used to determine the level of reliability and validity of the final scale are also presented thereafter.

Factor analysis (FA) is a technique that is widely used in psychometrics (Rust & Golombok, 2009) and in the most simple terms is employed to identify groups or clusters of variables within a large data set. More specifically, Field (2009) has suggested that FA has three main uses: 1) to understand the structure of a set of variables, 2) to construct a questionnaire to measure an underlying variable and 3) to reduce a data set to a more manageable size while retaining as much of the original information as possible.

The decision of how many factors to retain is a critical consideration of FA. Eigenvalues are commonly used and the criterion for assessing Eigenvalues is that important factors will have Eigenvalues greater than 1.0 (Kaiser Criterion). However, this generally leads to the inclusion of too many factors, especially when there are

many variables. Consequently, the use of 'Scree plots' (Cattell, 1966) and parallel analysis (Carraher & Buckley, 1995; Horn, 1965) can assist with this process (Cattell, 1966). The interpretation of the Scree plot is subjective; however, there is usually an evident change in the slope of the plot. Most commonly, the number of factors is deduced by selecting those that lie above the 'elbow' of the sloping line.

Parallel analysis is a further method that helps to guide this research decision making process (Carraher & Buckley, 1995; Horn, 1965). Hayton, Allen and Scarpello (2004) have upheld that parallel analysis is one of the most accurate factor retention methods despite also being one of the most underutilised. Factor analysis includes the calculation of Eigenvalues that measure the amount of variation in the total sample accounted for by each factor. The underlying rationale for parallel analysis is that the Eigenvalues of the salient factors from real data with a valid latent factor structure should be larger than the Eigenvalues of the corresponding factors that are generated from random data (Ford, MacCallum, & Tait, 1986; Lautenschlager, 1989). Similarly to the cognitive interviewing pre-test techniques implemented for a previous phase of the current research (Chapter III), parallel analysis is not always undertaken by scale development researchers. However, the method is receiving increasingly more attention in the literature as a method to support EFA (e.g., Liu & Rijmen, 2008).

The majority of factor analytic studies have applied the technique to help understand the correlation between variables. Factors represent underlying hypothetical constructs that can often be used to help understand and explain the data. For example, historically, trait theorists have used the technique extensively in an attempt to understand and measure personality (e.g., Eysenck, 1953). FA has been subject to debate regarding definitions, the mathematical basis and the appropriateness of its use (e.g., Cattell, 1978; Kline, 2000; Nunnally, 1978). This perhaps highlights the importance of the necessity for detailed and transparent reporting of the procedures undertaken during the research process in order to help refine sound reasoning for the methods employed and the most appropriate application. Despite debate, Kline (2000) maintained that FA makes it possible to simplify complex data and this is primarily conducted in one of three ways, or a combination of two or three. These are detailed below.

5.1.1 Exploratory Factor Analysis

EFA is used to map out the most important variables (Kline, 2000). The data is explored to identify a hidden structure (Rust & Golombok, 2009). EFA has been deemed as particularly appropriate for scale development or when there is little theoretical basis for specifying *a priori* the number and patterns of common factors (Hurley et al., 1997). However, Rust and Golombok (2009) have urged caution to a pure exploratory approach, instead proposing the use of theoretical presuppositions to assist with factor interpretations. Similarly to item analytical procedures, EFA is an iterative approach and a descriptive technique that provides several factor solutions to a set of data. The task of the researcher is to identify a factor structure that allows for the most plausible interpretation of the items. This requires a sound understanding of the constructs being examined that facilitates the research decisions made regarding factor structure. Therefore, if EFA is undertaken in this manner, the procedure does employ a degree of theoretical presuppositions as suggested by Rust and Golombok (2009). Indeed, the close association between factor analysis and construct validation has been noted repeatedly in the literature (e.g., Gorsuch, 1983; Guilford, 1946; Nunnally & Bernstein, 1994). However, an alternative method that utilises a much more theoretical and *a priori* knowledge approach is termed confirmatory factor analysis.

5.1.2 Confirmatory Factor Analysis

Confirmatory factor analysis (CFA) is used to confirm or support hypotheses (Kline, 2000). The method has built upon and replaced earlier procedures such as the Multitrait-Multimethod Matrix (MTMM Matrix) described by Campbell and Fiske (1959). Similarly to the IRT approach (see Chapter IV) the CFA method of analysis is employed when the intention is to categorize data into a proposed model with pre-determined, hypothesized groups of factor loadings often referred to as psychological dimensions. On occasion CFA will be used to confirm the results obtained from an EFA approach (Gerbing, & Hamilton, 1996). However, this often poses a challenge for smaller scale studies as two different data sets are required for each analysis. Therefore, twice the number of respondents are required to perform both analyses.

5.1.3 Principal Components Analysis

Principal Components Analysis (PCA) is commonly used as a method of exploratory data analysis and as such is often considered as a form of EFA. However, Kline (2000) has drawn some distinctions between factors and components. Kline (2000) proposed that components are real factors as they can be directly derived from the data of a study and that factors are hypothetical because they are estimated from the data. Component analysis is mathematically far more simple than factor analysis (Kline, 2000) and Harman (1976) has argued that when there are large matrices of variables there is little difference between the results of PCA and FA. Popular statistical software packages such as SPSS employ a PCA approach to EFA.

5.1.4 Factor Rotation

A final consideration of FA is the rotation of factors. There have been a number of suggestions as to how best to define the notion of rotation (e.g., Bryant & Yarnold, (1995; McDonald, 1985; Vogt, 1993). However, the primary function of rotation is to assist with the interpretation of the factor solution. For example, Yaremko, Harari, Harrison, and Lynn (1986) have suggested that the rotation of the factor axes (dimensions) identified in the initial extraction of factors are often rotated in order to obtain simple and interpretable factors.

Rotation methods are either orthogonal or oblique. Orthogonal rotation methods assume that the factors in the analysis are uncorrelated while oblique rotation methods assume that the factors are correlated (Gorsuch 1983). Typically, one of five rotation methods are used in FA and this is reflected by the options available to researchers who utilise the statistical software package SPSS. These rotation methods are known as varimax, direct oblimin, quartimax, equamax, and promax. Three of these are orthogonal (varimax, quartimax, and equimax), and two are oblique (direct oblimin and promax). Gorsuch (1983) has recommended rotating with varimax for orthogonal data or promax for oblique. Similarly, Kim and Mueller (1978) recommended that particularly for beginners, one of the more commonly available methods of rotation should be used such as varimax for orthogonal data and direct oblimin for oblique.

The focus of the current study is exploratory in nature (i.e., there is no theoretical underpinning of the items to be included in the analysis) and in this respect the literature indicated that an EFA approach to analysis was most appropriate. Furthermore, because PCA can be utilised with ease by using SPSS software and there is little difference between PCA and true factor analytical approaches (Harman,1976) a PCA approach was considered the most feasible. Additionally, to help guide the factor structure decision making process (Carraher et al.,1965) and to enhance accuracy (Hayton, et al., 2004) parallel analysis was also included in this phase of research. The literature further indicated that a varimax rotation was most suited to the orthogonal nature of the current data set (Gorsuch, 1983) and the most feasible (e.g., Kim & Mueller, 1978). The specific methods employed, including the procedure and analyses are detailed in the methods section (5.3-5.35).

5.1.5 Reliability, Validity and Acceptability

There are a range of methods available to scale researchers to assist in demonstrating the levels of reliability and validity of a measure. The concept of reliability has been identified as a fundamental way to reflect the amount of error inherent in any measurement (Streiner & Norman, 2008). While the concept of validity has been described as a process of determining what, if anything is being measured by the scale (Streiner & Norman, 2008).

Streiner and Norman (2008) have suggested that test-retest reliability is the usual approach to establishing the reliability of a scale that aims to measure self-rated reports of psychological function and this approach has also been described as the most straightforward (Rust & Golombok, 2009). This approach was utilised at the item level during the item analyses phase of scale construction (Chapter IV) and was used again to assess the reliability of the final scale for the current study. An assessment of the internal consistency of the scale under construction was also conducted at item level (Chapter IV) and was assessed again at scale level for the current study.

Selecting the most appropriate methods of assessing the validity of a new scale is an important part of scale construction. Streiner and Norman (2008) maintain that concurrent and construct validity analyses are most typical. Concurrent validity is

used when a new measure is being designed to replace or complement an existing scale that aims to measure the same construct (Rust & Golombok, 2009) and is therefore not relevant for the purpose of the current study. However, construct validity has been described as the primary form of validation underlying the trait-related approach to psychometrics where the construct being measured is not directly observable. Furthermore Cronbach and Meehl (1955) asserted that construct validity must be investigated whenever no criterion or universe of content is accepted as entirely adequate to define the quality to be measured. Construct validity refers to a wide range of approaches (Streiner & Norman, 2008) and the task of the researcher is to select those components of construct validity that are the most appropriate based upon the data and resources available.

In simple terms, the convergent component of construct validity investigates the degree to which constructs that should be related are statistically correlated and has been identified as an important part of the validity process by scale developers (e.g., Bowling, 1997). Schunemann et al. (2010) report that Pearson's correlation coefficients of less than 0.2 are very weak, from 0.2 to 0.35 are weak, from greater than 0.35 to 0.5 are moderate and of more than 0.5 are strong. Data were available for a measurement of respondents exercise-related mood (SEES; McAuley & Courneya, 1994) and generic QoL (WHOQOL-BREF; Skevington et al., 2004). Therefore, the convergent component of construct validity could be assessed and was included in the current study (see 5.3.3 for more detail). Data were also available regarding at what stage of referral respondents were at and therefore it was also possible to assess the known groups component of construct validity.

Content validity has been deemed as fundamental to psychometrics and is the main basis by which any test construction programme is judged (Rust & Golombok, 2009 p. 79). Content validity was a consideration that ran throughout the total phases of the current research. However, content validity was given particular attention in the final phase of scale development as the this stage of research determined the properties of the final measure. The items from the original 50-item pool that had been rejected were reviewed and the conceptual framework of exercise-related life-quality developed from focus groups (Chapter II) was also revisited from a content validity perspective.

If a scale is acceptable for use by the population for whom the measure is intended, it is logical that respondents will provide an answer to most of the items contained within it. Therefore, the acceptability of a measure is often considered by an analysis of the frequency of missing data and is sometimes coupled with cognitive pre-testing results if this phase of research is undertaken (e.g., Radloff, 1977). The acceptability of the scale under construction was also a consideration of the current and final phase of research.

5.2 Aims

The aims of the current research were:

- To understand the factor structure of the 29 items retained from item analyses and to reduce the data set further, rejecting those items that do not uniquely contribute to a factor within the factor structure while ensuring adequate breadth of measurement of the final scale and retaining as much of the original information as possible.
- To assess the reliability of the final scale structure through means of an assessment of internal consistency and test re-test reliability.
- To assess the validity of the final scale by means of an analysis of convergent and known groups components of construct validity.
- To assess the content validity of the final scale in view of rejected items and the conceptual framework of exercise-related life-quality that had been developed from focus groups data. To assess the overall acceptability of the measure and to re- organize the final set of scale items into a logical order that reflects the factor structure.

5.3 Methods

The methods utilised for this stage of scale construction and initial investigations of reliability and validity followed a series of psychometric analyses which included EFA, an assessment of reliability at scale level which included internal consistency (Cronbach's alpha) and test-retest reliability (intraclass) correlation, construct validity and an assessment of the acceptability of the final scale. Such procedures are

representative of what has been described as standard methodology for quality of life measure development (Schunemann et al., 2010) and the methods employed by the current study are also informed by a proposed framework to scale validation (Kirshner & Guyatt, 1985).

5.3.1 Participants

Participants were recruited in the exact same manner as for the item analyses (i.e., exercise referral instructors from each recruitment site approached individuals who were either at the start, in the middle or who had been exercising for up to 12 months post-referral to ask if they would volunteer to complete the test questionnaire). The same cohort (N = 278) was used for PCA. For the purposes of EFA, the size of the sample has been deemed as important in order to gain reliable factors (Kline, 2000) and a minimum of 200 participants has been suggested for a robust analysis (Guilford, 1956). However, there is some debate within the literature regarding adequate sample sizes for EFA and such debate is usually focussed upon suggestions regarding a minimum sample size or participants-to-item ratios. For example, Guadagnoli and Velicer (1988) maintain that absolute minimum sample sizes, rather than participant-to-item ratios, are more relevant. Recommendations for this minimum sample size approach range from 50 (Barrett & Kline, 1981) to 400 (Aleamoni, 1976). Alternatively, a minimum participant-to-item ratio of five-to-one has been suggested (Gorsuch, 1983; Hatcher 1994) and a higher ratio of ten-to-one (Nunnally, 1978).

Neither approach appears to be more effective than the other and it has been suggested that no single participant-to-item ratio approach will be suitable in all cases (MacCallum, Widaman, Preacher, & Hong, 2001). Perhaps one of the central considerations to scale developers is weighing-up the desired sample size against the feasibility of collecting this desired volume of data. This is often determined by the environment in which the data is collected and the inherent time and resource implications of doing so. For the purposes of the current study, a sample size of 278 fulfils the recommended minimum of 200 as suggested by Guilford (1956) and the minimum participant-to-item ratio of five-to-one (Gorsuch, 1983; Hatcher 1994). Although it is recognised that larger sample sizes are often better, it is also worth

noting that EFA is a purely exploratory procedure, useful for the initial phases of scale development and should typically precede CFA (Costello, & Osborne, 2005). Future development of the current scale includes CFA analyses (see Chapter VII) which will make use of a recommended sample size of 500 or more (Comfrey & Lee, 1992).

The statistical software package SPSS includes a Kaiser-Meyer-Olkin (KMO) statistic to ensure adequacy of the sample size. Values for the KMO statistic of 0.7-0.8 are good, 0.8-0.9 very good, and over 0.9 is considered excellent (Field, 2005). The sample size used in the current study had a Kaisser-Meyer-Olkin value of .863, supporting the adequacy of the sample size and Bartlett's Test of Sphericity reached statistical significance at the $p<.05$ level, thus demonstrating the factorability of the correlation matrix. The determinant value for the final solution (5.00) indicated that the variables were not too highly-inter-related - that multicollinearity and/or singularity were not a problem.

5.3.2 Materials

Item analyses reduced an initial pool of 50 test items to 29 items (Chapter IV) and these were included in the current analyses (see section 5.3.5). Of the 29 items included in the analyses, six items were reverse-scored. Nine items that failed to meet the statistical criteria for item analyses were included in the current analyses for content validity reasons (Streiner & Norman, 2008) thus enhancing the breadth of the measure Cattell, 1973; Klein, 2000). All 29 items including those that were retained for content validity reasons can be reviewed in Table 22.

Table 22. Items included in the principal components analysis (N = 29)

Item Number and Wording
*+Q1) To what extent can you afford leisure facilities?
Q4) How much do you feel that lifestyle factors (e.g., transport, time, childcare, poor health & the weather) affect your ability to be physically active?
+Q5) How much do you compare yourself to others while exercising?
Q9) How much does any injury you may have prevent you from being physically active?
+Q10) How much do you currently enjoy physical activity?
+Q11) How much is social contact with others a part of your current lifestyle?
+Q14) How well do you feel you maintain eating habits that are beneficial to your health and any illness you may have?
Q17) How confident are you using gym equipment?
Q18) How confident are you that you can exercise in a leisure centre with minimum support?
Q22) How confident are you in your own physical ability?
Q23) How confident are you to join in activities with family and friends?
Q28) In general, how much opportunity do you have to make time for yourself?
Q29) How often do you experience physical pain?
Q30) How often do you feel you have to manage any symptoms of pain?
Q31) How often do you feel you have to manage the symptoms of any illness you have?
Q32) How often do you feel you have achieved something in terms of your physical fitness?
Q33) How would you rate your energy levels?
+Q34) How would you rate your outlook on life?
Q35) How would you rate your level of self-esteem?
Q36) How would you rate your physical health?
+Q38) How would you rate your sense of well-being?
Q41) How would you rate the quality of your sleep?
Q42) How would you rate your memory function?
Q44) How would you rate your ability to manage stress?
Q45) How would you rate your ability to manage your mood?
Q46) How would you rate your ability to manage your weight?
+Q47) How would you rate your current life expectancy?
Q48) How would you rate your ability to undertake everyday tasks that require some level of physical activity?
+Q49) How would you rate your overall quality of life?

Note. * Denotes and item re-worded following item analyses + denotes an item retained from item analyses on the grounds of content validity.

5.3.3 Selection of Construct Measures and Correlation Hypotheses

The rationale for selecting the WHOQOL-BREF (Skevington et al., 2004) as the generic QoL measure with which to conduct analyses of convergent validity was based upon three key factors: questionnaire performance (psychometric properties), questionnaire length and appropriateness of questions for exercise referral populations.

The WHOQOL-BREF (Skevington et al., 2004) is a 26-item version of the WHOQOL-100 assessment and consists of the measurement of four domains: physical, psychological, social and environmental. The psychometric properties were analysed using cross-sectional data obtained from a survey of adults carried out in 23 countries (N = 11,830). Analyses of internal consistency, item–total correlations, discriminant validity and construct validity through CFA, indicated that the WHOQOL-BREF had good to excellent psychometric properties of reliability and performed well in preliminary tests of validity.

The questionnaire under development had 50 original test-items in total (only 29 of which were used in the current analyses) therefore, it was particularly important to select a comparative measure that not only performed well, but that was easy to administer and did not impose unnecessary burden upon the respondent, the importance of which has been raised previously (Skevington et al., 2004). The 26 items of the WHOQOL-BREF (Skevington et al., 2004) were sufficient to allow for comparative statistical analyses while minimising respondent burden.

It was expected that those individuals who report higher perceptions of exercise-related life-quality would also perceive positive perceptions of generic QoL. Specifically, it was predicted that there would be a strong positive correlation between the test measure total scores and the physical health and also psychological health dimensions of the WHOQOL-BREF because the items included in the scale are most representative of these constructs. It was also predicted that there would be a moderate positive correlation between the test measure total scores and social relationships and the environment dimensions of the WHOQOL-BREF because the items included in the measure do reflect these constructs but not as thoroughly as

psychological and physical health. These analyses were used to contribute to the assessment of convergent validity. Correlations between the total scores of the test measure and the WHOQOL-BREF were not conducted as the WHOQOL-BREF is not scored as a single global score.

Perceptions of mood and mood management are especially relevant for clinical exercise populations (e.g., Lehrner et al., 1999; Mino, & Hicks, 2004; Petajan, Gappmaier, White, Spencer; Stewart et al., 2003). This documented relationship was the reasoning behind the selection of a measure of mood for the purposes of convergent validity analyses. Again, it was important that the measure had suitable psychometric properties, was not too lengthy to avoid any unnecessarily increase in respondent burden and that the measure was relevant for exercise settings.

The SEES (McAuley & Courneya, 1994) is a measure designed to assess the global psychological responses to the stimulus properties of exercise and as such was selected as an appropriate measure to assess the convergent validity of the current scale under construction. The SEES consists of a three-factor measure. Two of the three factors correspond with the positive and negative poles associated with psychological health, positive well-being and psychological distress, while the third factor represents subjective indicators of fatigue. Perhaps one of the limitations of the initial EFA methods of analysis to determine this factor structure was that undergraduate university students (N = 454; age M = 20.78 years, SD ± 2.18) were used to collect the required data. Thus the transferability of the results to more clinical settings, which is where a measure of subjective exercise experience would be particularly suited, was questionable at this stage of the scale's development. However, in further studies, McAuley and Courneya (1994) utilised CFA with 51 middle aged (M = 55 years) males (n = 27) and females (n = 24) who were enrolled in a programme of exercise similar to that of exercise referral. The results supported the initial EFA findings and although the authors have encouraged further exploration of the psychometric structure of the SEES (McAuley & Courneya, 1994), the short 12-item nature of the measure coupled with the relevance and appropriateness of the population that was used to generate the data for CFA meant that the measure was particularly suited to the requirements of convergent validation analyses.

With respect to mood, for the same reasoning as the strength of relationships predicted for the WHOQOL-BREF, it was expected that total scores from the test measure would have a strong positive correlation with the positive well-being dimension of the SEES and that there would be a strong negative correlation with the total scores from the test measure and the psychological distress dimension of the SEES. It was also expected that the total scores from the test measure would have a moderate to weak negative correlation with the fatigue dimension of the SEES. These analyses were also used for the purposes of convergent validity. Correlations between the total score of the test measure and that of the SEES were not conducted as the SEES is not scored as a single global score.

Additionally, an analysis of the test measure's sub-scales that were identified by EFA in the current study was undertaken with respect to the domains of the WHOQOL-BREF and the SEES. Similar to those predictions of test measure total score correlations, it was predicted that the strongest positive correlations would be observed with dimensions of the test scale that most closely represented positive mental well-being and those of the WHOQOL-BREF. For example, between those items that were represented by the mental and physical well-being component of the conceptual framework and the psychological and physical domains of the WHOQOL-BREF. The same prediction was made for correlations between these test items and the psychological well-being dimension of the SEES. Again, the strongest negative correlations were predicted between the test measure sub-scales and those domains contained within the WHOQOL-BREF and SEES that are likely to decrease as exercise-related life-quality perceptions increase. For example, it was expected that those items currently grouped together as psychological well-being or those grouped as confidence and motivation would show the strongest negative correlations with the psychological distress dimension of the SEES. In terms of the WHOQOL-BREF because all the domains are positively constructed it was predicted that items conceptually grouped as discomfort under the discomfort and fitness conceptual framework would have the strongest negative correlations with those items contained within the physical dimension of the WHOQOL-BREF.

5.3.4 Procedure

The data used for the current analyses (see section 5.3.5) was the same set that had been used for the previously conducted item analyses and as such the data had been cleaned of any significant chunks of missing data (whereby an entire page of questions had been missed, for example) and these questionnaires were not included in the analysis. The confidentiality of participants had been assured by assigning a code to their responses (see Chapter IV). These procedures carried forward into the PCA.

Initially, the 29 items were subject to PCA and the importance of each factor and how much variation was accounted for by each factor was assessed by Eigenvalues. The Scree plot generated from this analysis (Figure 13) was examined to further inform the likely factor structure and additional PCA was conducted which involved the systematic deleting of items from the item pool until a statistically and conceptually robust solution was reached.

Parallel analysis was used to guide decisions regarding the factor solution. The internal consistency of these items was also analysed using Cronbach's alpha, test re-test reliability was assessed at scale level, the convergent component of construct validity was assessed using Pearson's correlation, content validity was assessed by reviewing those items that were rejected from the original 50-item pool and also in view of the conceptual framework of exercise-related life-quality that had been developed. The acceptability of the final scale was assessed by an analysis of the frequency of missing data. This data is combined with cognitive pre-testing results (Chapter III) to include a more complete assessment of acceptability in Chapter VI. An additional FA, this time forcing a single factor solution was also undertaken to assess if the reduced item pool could be scored as a single global score to assist with the ease with which the measure could be used in practice.

The domain titles allocated to each of the resultant factors were given careful consideration and was guided by the recommendations of Pett, et al. (2003). A series of peer de-briefing sessions (Spall & Stephen, 1998) were also undertaken to reach the most appropriate descriptors that not only represented the item content of each factor but that would be conceptually meaningful to both patients and practitioners. Once the factor structure and domain titles had been established the remaining items

were re-organized to follow a logical sequence, respecting the guidelines that were followed during the initial construction of items (Hague, 1993) and grouped together with other items that reflected their respective factor.

5.3.5 Analyses

Principal component and parallel analysis were conducted using SPSS (v17.0) and based upon the recommendations of Gorsuch (1983) and Kim and Mueller (1978) a varimax rotation was used to assist with the interpretation of the final factor structure. Field (2009) has advised that a factor loading of >0.3 should be considered significant. However, Stevens (2002) has also made recommendations based upon sample size. The smaller the sample size, the greater the factor loading required. For a sample size of 200, Stevens (2002) proposed a critical value of >0.367 against which factor loadings can be compared. Although, Costello and Osborne (2005) have argued that a minimum level of 0.4 should be considered a best practice approach. In the current study, a critical value of >0.4 was used to honour this best practice recommendation.

A parallel analysis was conducted to explore the level of confidence that could be attributed to the factor structure produced by PCA and the internal consistency (Cronbach's alpha) and test-retest reliability (intraclass correlation co-efficient) of the final item set was also analysed using SPSS (v17.0). Following this the convergent component of construct validity was assessed by an analysis of the correlations (Pearson's correlation) between scores from the test measure and those generated from the dimensions of the WHOQOL-BREF and SEES. Based upon the methods employed by previous QoL scale development research, the current study considered Pearson's correlation coefficients of less than 0.2 as very weak, from 0.2 to 0.35 as weak, from greater than 0.35 to 0.5 as moderate and of more than 0.5 as strong (Schunemann et al., 2010). The known groups component of construct validity was assessed by a three-way ANOVA that assessed the level of total test score variance for those at the start, midway and who had completed their programme of exercise. Again, data were analysed using SPSS (v17.0). Data handling procedures replicated those that were undertaken for the purposes of item

analyses (Chapter IV) and the process of allocating the titles given to each of the three factors was guided by the recommendations of Pett, et al. (2003).

5.4 Results

The results from each of the analytical procedures undertaken for the current study are reported below.

5.4.1 Principal Components Analysis

PCA of 29 items resulted in seven items being rejected from the item-pool and the remaining 22 items were able to be grouped into a statistically and conceptually coherent three-factor structure. An analysis of Eigenvalues initially offered a seven factor solution, however factors four, five six and seven contributed very little to the percentage of total variance within this seven-factor structure. The values were factor 7 (3.62%), factor 6 (3.82%), factor 5 (4.74%) and factor 4 (4.86%) respectively. Furthermore, an analysis of the Scree plot suggested a three-factor structure (Figure 13) and as a consequence the PCA analysis was re-run, instead forcing a three-factor structure and items were eliminated if they failed to load at the critical value (>0.4) or if they cross-loaded meaning that an item loaded at the critical value (>0.4) onto more than a single factor. This careful and systematic process resulted in a 22-item three-factor solution. Table 23 documents the factor loadings of the resultant three-factor solution with domain titles.

Figure 13. Scree plot indicating a three-factor solution

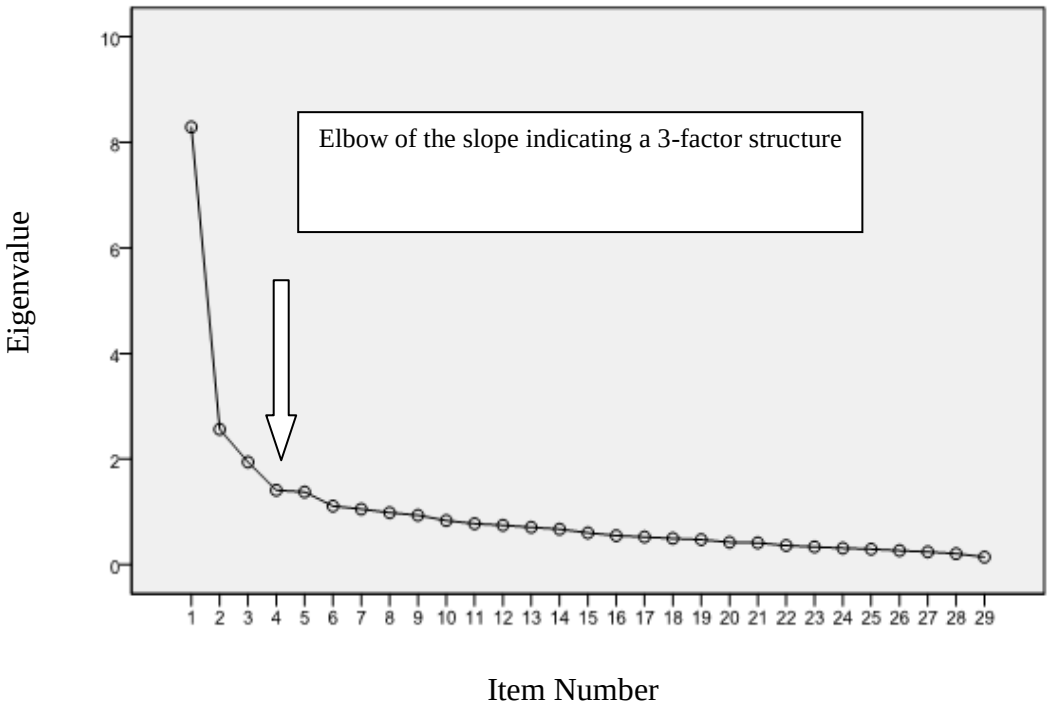


Table 23. Final 22-item three-factor solution and domain titles

Item	Mental and Physical Well-being	Injury, Pain and Illness	Physical Activity Facilitators
Q9 Injury Prevention	.12	**.59**	.27
Q10 Enjoyment of PA	.29	.06	**.53**
Q17 Gym Equipment Confidence	-.07	.14	**.79**
Q18 Minimum Support	-.02	.17	**.74**
Q22 Confidence Physical Ability	.35	.27	**.63**
Q23 Confidence Family & Friends	.39	.20	**.49**
Q28 Time for Self	.23	-.07	**.47**
Q29 Frequency of Pain Experience	.03	**.89**	.04
Q30 Frequency of Pain Symptom Management	.07	**.89**	.07
Q31 Illness Symptom Frequency Management	.20	**.69**	-.07
Q32 Physical Fitness Achievement	.31	-.12	**.54**
Q34 Outlook on Life	**.74**	.09	.15
Q35 Self Esteem	**.64**	.09	.30
Q36 Physical Health	**.51**	.48	.28
Q38 Sense of Well-Being	**.76**	.23	.14
Q41 Sleep Quality	**.47**	.24	-.03
Q42 Memory Function	**.48**	-.00	.15
Q44 Stress Management	**.77**	-.01	.13
Q45 Mood Management	**.77**	-.04	.07
Q46 Weight Management	**.44**	.21	.24
Q47 Life Expectancy	**.60**	.11	.25
Q49 Overall QoL	**.68**	.31	.13

Note. Items in bold denote factors that load at the critical value (>0.4). Values in bold denote Eigenvalues greater than parallel analysis values.

The three factor solution explained a total of 50.68 percent of the variance, with factor one (mental and physical well-being) contributing 31.48 percent, factor two (injury pain and illness) contributing 10.68 percent and factor three (physical activity facilitators) contributing 8.52 percent of the variance respectively. The item-total correlations for each item included in the final 22-item three-factor solution met the statistical criteria of >0.3 (Skevington et al., 2004).

5.4.2 Parallel Analysis

A parallel analysis of the final 22-item three-factor solution revealed that the Eigenvalues of the first three components were greater than the parallel analysis equivalent and thus confirmed that all three factors were robust 'true' factors and not just a random aggregation of items (Table 24). Factors four and five had Eigenvalues of 1.16 and 1.05 respectively. It is likely that these are spurious and represent fairly random aggregations of items. It is reasonable therefore to ignore them even though they have Eigenvalues of greater than one.

Table 24. Eigen and parallel analysis values of the three-factor solution (N = 22)

Component	Eigenvalue	Parallel Analysis Value
1	**6.93**	1.54
2	**2.35**	1.45
3	**1.87**	1.38
4	1.16	1.32
5	1.05	1.26
6	.93	1.21
7	.88	1.17
8	.85	1.12
9	.72	1.08
10	.71	1.04
11	.58	1.00
12	.54	.96
13	.53	.92
14	.47	.88
15	.45	.85
16	.41	.81
17	.36	.77
18	.31	.73
19	.28	.70
20	.27	.66
21	.22	.61
22	.14	.56

5.4.3 Content Validity

During focus group and item analytical procedures a conceptual framework was developed to guide item retention and deletion. This conceptual framework comprised the following: leisure facilities and lifestyle PA (12 items), confidence and motivation (12 items), discomfort and fitness (4 items) and mental and physical well-being (22 items). The resultant final factor titles were mental and physical well-being (11 items), injury pain and illness (4 items) and physical activity facilitators (7 items).

Of the nine items that during item analyses were retained for content validity reasons and were included in the PCA item pool (N = 29); five were retained and included in the final 22-item three-factor solution. These items were Q10 regarding the

enjoyment of PA, Q34 regarding perceived outlook on life, Q38 regarding perceived sense of well-being, Q47 regarding perceived life expectancy and Q49 regarding perceived overall life-quality. A total of four reverse-scored items were included in the final solution and these items all grouped together into factor two; injury, pain and illness. Table 25 illustrates those items rejected from the original 50-item test pool and Table 26 illustrates the final 22 items and the factor and conceptual component of life-quality to which they belong.

Table 25. Items rejected from the original 50-item test pool

Item
Q1 To what extent can you afford leisure facilities
Q2) To what extent do you feel you have a choice regarding the type of exercise you undertake?
Q3) To what extent do you feel you have a choice regarding the time at which you exercise?
Q4) How much do you feel that lifestyle factors (e.g., transport, time, childcare, poor health & the weather) affect your ability to be physically active?
Q5) How much do you compare yourself to others while exercising?
Q6) To what extent do you worry about exercising in an environment with others?
Q7) To what extent do you worry about participating in structured exercise?
Q8) To what extent do you worry about the prospect of exercising in a fitness suite/gym?
Q11) How much is social contact with others a part of your current lifestyle?
Q12) How important is it for you to manage your weight?
Q13) How much do you feel that you incorporate physical activity into your daily lifestyle?
Q14) How well do you feel you maintain eating habits that are beneficial to your health and any illness you may have?
Q15) In terms of exercise how competitive are you?
Q16) In terms of exercise how determined are you?
Q19) How confident are you in your ability to participate in regular physical activity and exercise?
Q20) How confident are you in your ability to undertake daily tasks that do not necessarily involve physical activity?
Q21) In general, how confident are you around other people?
Q24) How important is it to you right now to maintain a physically active lifestyle?
Q25) How confident are you, right now that you can maintain a physically active lifestyle?
Q26) How motivated are you to maintain a physically active lifestyle?
Q27) How likely are you to maintain a physically active lifestyle?
Q33) How would you rate your energy levels?
Q37) How would you rate your mental health?
Q39) How would you rate the quality of support you receive from leisure staff to be physically active?
Q40) How would you rate the quality of support you receive from family and friends to be physically active?
Q43) How would you rate your level of motivation to be physically active?
Q48) How would you rate your ability to undertake everyday tasks that require some level of physical activity?
Q50) How would you rate your current knowledge of the benefits of physical activity and exercise for health?

Table 26. The factor and conceptual framework component of final items

Item	Factor	Conceptual Component
Q9 Injury Prevention	Injury Pain & Illness	Leisure Facilities & LPA
Q10 Enjoyment of PA	PA Facilitators	Leisure Facilities & LPA
Q17 Gym Equipment Confidence	PA Facilitators	Confidence & Motivation
Q18 Minimum Support	PA Facilitators	Confidence & Motivation
Q22 Confidence Physical Ability	PA Facilitators	Confidence & Motivation
Q23 Confidence Family & Friends	PA Facilitators	Confidence & Motivation
Q28 Time for Self	PA Facilitators	Leisure Facilities & LPA
Q29 Frequency of Pain Experience	Injury Pain & Illness	Discomfort & Fitness
Q30 Frequency of Pain Symptom Management	Injury Pain & Illness	Discomfort & Fitness
Q31 Illness Symptom Frequency Management	Injury Pain & Illness	Discomfort & Fitness
Q32 Physical Fitness Achievement	PA Facilitators	Discomfort & Fitness
Q34 Outlook on Life	Mental & PWB	Mental & PWB
Q35 Self Esteem	Mental & PWB	Mental & PWB
Q36 Physical Health	Mental & PWB	Mental & PWB
Q38 Sense of Well-Being	Mental & PWB	Mental & PWB
Q41 Sleep Quality	Mental & PWB	Mental & PWB
Q42 Memory Function	Mental & PWB	Mental & PWB
Q44 Stress Management	Mental & PWB	Mental & PWB
Q45 Mood Management	Mental & PWB	Mental & PWB
Q46 Weight Management	Mental & PWB	Mental & PWB
Q47 Life Expectancy	Mental & PWB	Mental & PWB
Q49 Overall QoL	Mental & PWB	Mental & PWB

Note. LPA represents lifestyle physical activity. PWB represents physical well-being.

The 22 items included in the three-factor solution also loaded onto a single factor (>0.4). As a consequence, the final scale may be scored as three separate dimensions of mental and physical well-being, injury, pain and illness and PA facilitators or as a single global score. Table 27 documents the factor loadings of the single factor solution.

Table 27. Factor loadings of a single factor solution (N = 22)

Item	Factor Loading
Q9 Injury Prevention	.46
Q10 Enjoyment of PA	.51
Q17 Gym Equipment Confidence	.39
Q18 Minimum Support	.41
Q22 Confidence Physical Ability	.69
Q23 Confidence Family & Friends	.62
Q28 Time for Self	.38
Q29 Frequency of Pain Experience	.40
Q30 Frequency of Pain Symptom Management	.47
Q31 Illness Symptom Frequency Management	.40
Q32 Physical Fitness Achievement	.45
Q34 Outlook on Life	.68
Q35 Self Esteem	.67
Q36 Physical Health	.72
Q38 Sense of Well-Being	.75
Q41 Sleep Quality	.44
Q42 Memory Function	.45
Q44 Stress Management	.65
Q45 Mood Management	.62
Q46 Weight Management	.55
Q47 Life Expectancy	.63
Q49 Overall QoL	.72

Note. Factors that load at the critical value (>0.4).

5.4.4 Reliability

The internal consistency of all 22-items included in the three-factor solution was good with a Cronbach's alpha of 0.88 and alpha could not be improved by deleting any of the 22 items included in the final scale. Additionally, all 22-items met the item-total correlation criteria of >0.3 (Skevington et al., 2004) (Table 28).

The internal consistency (Cronbach's alpha) of the items that comprise the mental and physical well-being dimension was 0.87. For injury pain and illness 0.81 and for physical activity facilitators 0.77. Test re-test (intraclass correlation) at scale level indicated a significant level of reliability ($r = .72$; $p < .001$).

The test-retest reliability (intraclass correlation) of the scale at sub-scale level was ($r = .95$; $p < .001$) for mental and physical well-being, ($r = .80$; $p < .001$) for injury pain and illness and ($r = .86$; $p < .01$) for physical activity facilitators.

Table 28. Item-total correlations and Cronbach's alpha values of the 22-item three-factor solution

Item	Item-Total Correlation	Cronbach's Alpha If Item Deleted
Q9 Injury Prevention	.46	.88
Q10 Enjoyment of PA	.41	.88
Q17 Gym Equipment Confidence	.31	.88
Q18 Minimum Support	.37	.88
Q22 Confidence Physical Ability	.60	.87
Q23 Confidence Family & Friends	.54	.87
Q28 Time for Self	.32	.88
Q29 Frequency of Pain Experience	.40	.88
Q30 Frequency of Pain Symptom Management	.49	.87
Q31 Illness Symptom Frequency Management	.36	.88
Q32 Physical Fitness Achievement	.36	.88
Q34 Outlook on Life	.567	.87
Q35 Self Esteem	.58	.87
Q36 Physical Health	.67	.87
Q38 Sense of Well-Being	.68	.87
Q41 Sleep Quality	.37	.88
Q42 Memory Function	.38	.88
Q44 Stress Management	.52	.87
Q45 Mood Management	.49	.87
Q46 Weight Management	.45	.88
Q47 Life Expectancy	.55	.87
Q49 Overall QoL	.65	.87

5.4.5 Construct Validity

The descriptive statistics of the data from the test score total score and the total scores from each of the dimensions of the SEES and WHOQOL-BREF can be reviewed in Table 29.

Table 29. Descriptive statistics of the test measure total score and domain scores of the SEES and WHOQOL-BREF

	Mean	SD	Range (Min/Max)
Test Measure	75.02	8.92	40.00 – 97.00
SEES PWB	18.34	5.23	4.00 – 28.00
SEES PD	6.88	4.39	4.00 – 25.00
SEES FAT	10.42	6.13	4.00 – 28.00
QOLBREF PHYS	23.10	3.18	13.00 – 32.00
QOLBREF PSYCH	20.56	3.12	11.00 – 27.00
QOLBREF SOC	11.1	2.49	3.00 – 11.00
QOLBREF ENV	31.25	4.84	9.00 – 40.00

Note. Psychological well-being (PWB), psychological distress (PD), fatigue (FAT) physical (PHYS), psychological (PSYCH), social (SOC) and environment (ENV).

As expected there were strong correlations observed between the test-measure total score and the dimensions of physical health ($r = .64$; $p < 0.01$) and psychological health ($r = .62$; $p < .01$) of the WHOQOL-BREF. Moderate correlations were observed between the test-measure total score and social relationships ($r = .47$; $p < 0.01$) and environment ($r = .56$; $p < 0.01$). These data support the convergent validity of the test measure and because of the variance in the strength of the correlations between the total test-measure scores and the respective dimensions of the WHOQOL-BREF, the data would also suggest that the test instrument is measuring components of life-quality that differ to those generic dimensions assessed by the WHOQOL-BREF (Streiner & Norman, 2008).

Additional support for the convergent validity of the test measure was offered from the observed strong correlation with the test-measure total score and those from the positive well-being dimension of the SEES (r = .55; p <0.01). This was further supported by observing moderate negative correlations with the psychological distress dimension of the SEES (r = -.45; p <0.01) and a weak positive correlation with the fatigue (r =.30; p <0.01).

The convergent validity of the scale was also assessed by an examination of the scores obtained from each of the respective domains of the new measure and those of the WHOQOL-BREF and the SEES. The descriptive statistics of the data used for this analyses can be found in Table 30 and the correlation analyses can be reviewed in Table 31.

Table 30. Descriptive statistics for total dimension scores of the test measure WHOQOL-BREF and SEES

	Mean	SD	Range (Min/Max)
Test MPW	73.72	6.23	53.00 – 88.00
Test IPI	10.95	3.65	4.00 – 20.00
Test PAF	24.60	4.29	10.00 – 35.00
SEES PWB	18.34	5.23	4.00 – 28.00
SEES PD	6.88	4.39	4.00 – 25.00
SEES FAT	10.42	6.13	4.00 – 28.00
QOLBREF PHYS	23.10	3.18	13.00 – 32.00
QOLBREF PSYCH	20.56	3.12	11.00 – 27.00
QOLBREF SOC	11.1	2.49	3.00 – 11.00
QOLBREF ENV	31.25	4.84	9.00 – 40.00

Note. Mental and physical wellbeing (MPW), Injury pain and illness (IPI), physical activity facilitators (PAF), Psychological well-being (PWB), psychological distress (PD), fatigue (FAT) physical (PHYS), psychological (PSYCH), social (SOC) and environment (ENV).

Table 31. Sub-domain correlations between the test measure, WHOQOL-BREF and SEES

	Test Measure		
	MPW	IPI	PAF
SEES PWB	**.61**	-.27	.42
SEES PD	**-.52**	.34	-.30
SEES FAT	-.39	.37	-.28
QOLBREF PHYS	**.65**	**-.71**	**.52**
QOLBREF PSYCH	**.73**	-.35	.47
QOLBREF SOC	**.56**	-.24	.34
QOLBREF ENV	**.64**	-.38	.47

Note. Mental and physical wellbeing (MPW), Injury pain and illness (IPI), physical activity facilitators (PAF), Psychological well-being (PWB), psychological distress (PD), fatigue (FAT) physical (PHYS), psychological (PSYCH), social (SOC) and environment (ENV). Items in bold denote the strongest correlations. Values marked in bold denote the strongest correlations (>0.5).

Table 31 indicates that the correlations between each of the three dimensions of the new measure and those of the WHOQOL-BREF and SEES were significant according to the values suggested by Schunemann et al. (2010). The strongest positive correlations were observed between the mental and physical well-being dimension of the new measure and the psychological well-being dimension of the SEES ($r = .61$; $p < .01$), and the physical ($r = .65$; $p < .01$), psychological ($r = .73$; $p < 0.01$), social ($r = .56$; $p < .01$) and environmental $r = .64$; $p < .01$) dimensions of the WHOQOL-BREF. The strongest correlation between the physical activity facilitators domain of the new measure was with the physical dimension of the WHOQOL-BREF ($r = .52$; $p < .01$). The strongest negative correlations were observed between the injury pain and illness domain of the new measure and the physical dimension of the WHOQOL-BREF ($r = -.71$; $p < .01$) and the mental and physical well-being domain of the new measure and the personal distress domain of the SEES ($r = -.52$; $p < .01$).

The descriptive statistics of the data used to assess the known group validity component of construct validity can be found in table 32.

Table 32. Descriptive statistics for known group validity analyses

	Mean	SD	Range (Min/Max)
Programme Start	72.80	10.97	50.00 – 93.00
Mid Programme	77.76	11.23	45.00 – 105.00
Programme Completion	79.59	11.05	53.00 – 102.00

There were significant differences between the mean scores for each of the three groups at the <.05 level $F(2,226) = 5.07$, $p = .007$. Levine's statistic indicated that the variance between mean scores for the three groups was not significant ($p = .969$; $\underline{p} < .05$). The sample sizes were unequal (programme start, n = 45; mid programme, n = 123; programme completion n = 61) therefore, a Games-Howell procedure was used to assess the specificity of this variance. Data indicated that the largest significant difference in mean total scores was between those that had just started their exercise referral programme and those that had completed ($p = .006$; $\underline{p} < .05$). There was also a significant difference in mean total scores for those that had just started their referral programme and those that were midway through ($p = .03$; $\underline{p} < .05$). Differences in mean total scores for those respondents midway through their referral programme and those that had completed were also observed, although these differences were not significant ($p = .55$; $\underline{p} < .05$).

5.4.6 Acceptability

For the test measure, data were missing most frequently in response to levels of perceived confidence regarding the use of gym equipment (n = 8) (3%).

Items that had the lowest level of missing data included those regarding outlook on life (n = 1), self-esteem (n = 1) and sleep quality (n = 1). The item regarding overall QoL was responded to by all 278 participants (100%).

The most frequent missing data for the SEES was in response to the item rating of how 'crummy' the respondent felt (n = 11) (4%). No single item contained within the SEES received a 100% response rate. However, the lowest amount of missing data

were in response to the item requesting a rating of how ‘great’ respondents felt (n = 2).

The most frequent missing data for the WHOQOL-BREF was in response to the sex life satisfaction item (n = 42) (15%). Nine of the items that comprise the WHOQOL-BREF received 100% response rate (N = 278). These were items regarding QoL, health satisfaction, concentration, energy level, money, leisure opportunities, support from friends satisfaction, health service satisfaction and frequency of negative feelings.

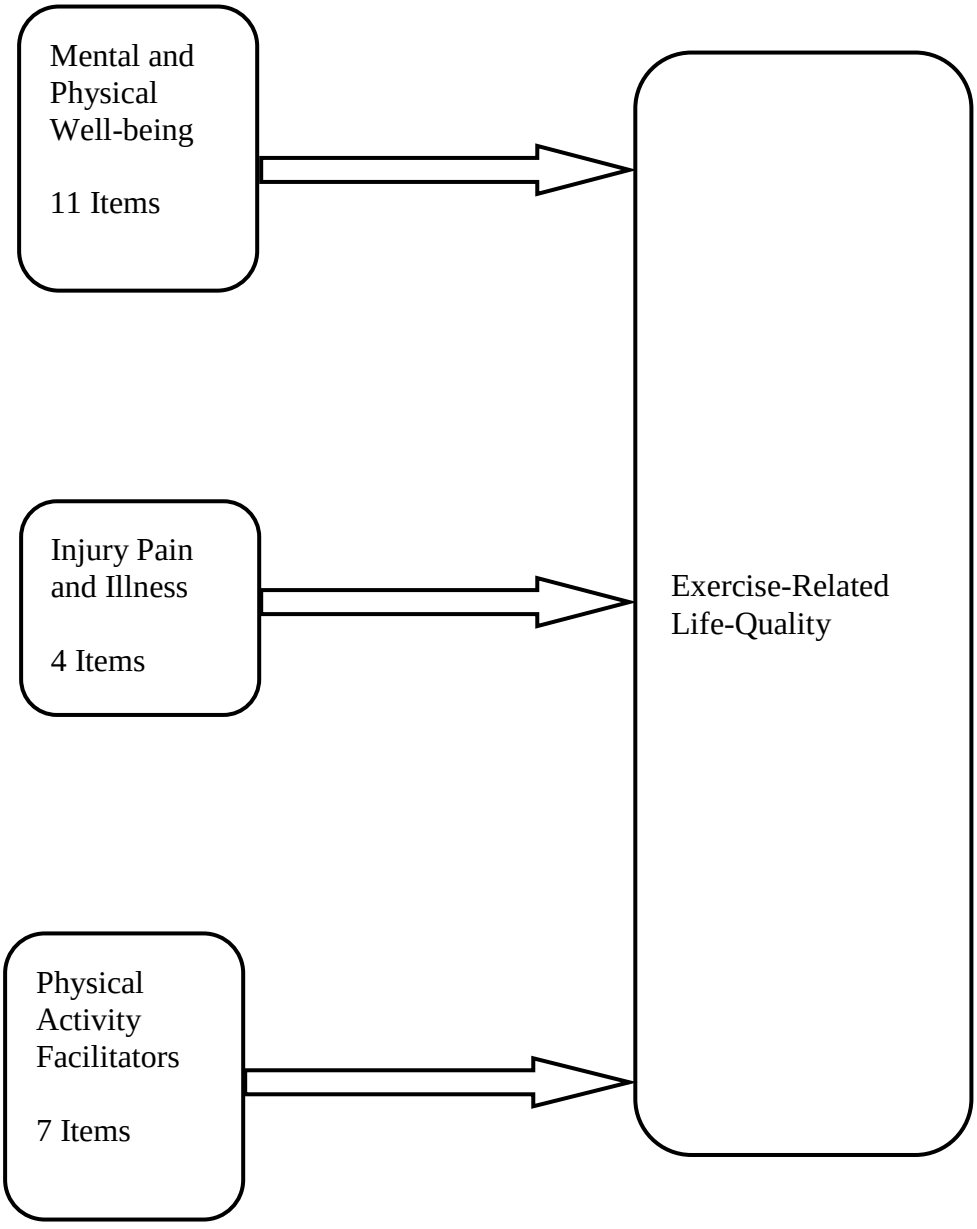
5.4.7 Organization of Items in the Final Draft Scale

Once the final 22-items had been grouped into their respective factors, the question numbers that had been allocated to these items in the original pool (i.e., 1-50) became redundant at this stage and new question numbers had to be allocated to the retained items in order to follow a numeric sequence (i.e., 1-22). In keeping with the guidelines and recommendations followed in the initial phases of item construction and questionnaire design (e.g., Hague, 1993), the final set of items were re-numbered so as to follow a numeric sequence and also in a manner that allowed items from the same factor to be listed together (Table 33). The reverse-scored items from the factor injury pain and illness were grouped halfway through the sequence of 22-items; following the factor mental and physical health and prior to physical activity facilitators because this was considered the most effective positioning of questions that are designed to prompt careful thinking before responding and detract respondents from entering into a pattern of responses. Figure 14 provides a visual representation of the three-factor structure with domain titles and the number of items representing each domain.

Table 33. Question order of final items

Item
Q1) How would you rate your outlook on life?
Q2) How would you rate your level of self-esteem?
Q3) How would you rate your physical health?
Q4) How would you rate your sense of well-being?
Q5) How would you rate the quality of your sleep?
Q6) How would you rate your memory function?
Q7) How would you rate your ability to manage stress?
Q8) How would you rate your ability to manage your mood?
Q9) How would you rate your ability to manage your weight?
Q10) How would you rate your current life expectancy?
Q11) How would you rate your overall quality of life?
Q12) How much does any injury you may have prevent you from being physically active?
Q13) How often do you experience physical pain?
Q14) How often do you feel you have to manage any symptoms of pain?
Q15) How often do you feel you have to manage the symptoms of any illness you have?
Q16) How much do you currently enjoy physical activity?
Q17) How confident are you using gym equipment?
Q18) How confident are you that you can exercise in a leisure centre with minimum support?
Q19) How confident are you in your own physical ability?
Q20) How confident are you to join in activities with family and friends?
Q21) In general, how much opportunity do you have to make time for yourself?
Q22) How often do you feel you have achieved something in terms of your physical fitness?

Figure 14. A three factor representation of exercise-related life-quality



5.5 Discussion

The primary purpose of the current phase of study was to reduce the test-item pool even further, establish the factor structure of the scale under construction and to assess the reliability, validity and acceptability of the final item pool.

A PCA approach to EFA reduced a 29-item pool to 22 items that were statistically robust enough to comprise the final exercise-related life-quality scale. Parallel analysis supported the resultant 22-item three-factor structure as the Eigenvalues of the three factors were all consistently higher than those produced by parallel analysis. Similarly, all item-total correlations met the statistical criteria of >0.3 (Skevington et al., 2004) and the internal consistency of the final scale as determined by Cronbach's alpha was good (0.88). Test re-test reliability (intraclass correlation) at scale level indicated a significant level of reliability ($r = .72$; $p < .001$).

The 22-item solution demonstrates a good range of items within three domains of exercise-related life-quality that represent those items of importance raised by the focus group participants well. This is evidenced because each category contained within the conceptual framework of exercise-related life-quality is represented by items included in the final scale. Questions regarding outlook on life, self esteem, physical health, sense of well-being, sleep quality, memory function, stress management, mood management, weight management, life expectancy and overall QoL (n = 11) comprise the mental and physical well-being domain. Questions regarding injury prevention, frequency of pain, frequency of pain symptom management and frequency of illness symptom management (n = 4) comprise the injury, pain and illness domain and questions regarding the enjoyment of PA, confidence using gym equipment, exercising with minimum support, confidence in physical ability, confidence around family and friends, time for self and frequency of physical fitness achievement (n = 7) comprise the physical activity facilitators domain. Coupled with the sub-domain convergent validity analysis, the items contained within each domain can provide insights into what is actually measured by the final scale. This is addressed a little later in the current section and the broader implications for QoL assessment are addressed in Chapter VII.

In more general terms, the item content of each of these domains is such that the final scale is likely to represent a measure of exercise-related life quality that is neither too

specific nor too broad. The mental and physical well-being domain captures some of those more generic aspects of life-quality found in other measures. For example, questions regarding sleep quality are found in the physical domain of the WHOQOL-BREF (Skevington et al., 2004) and self-esteem are also found in the psychological domain of the same measure. The current research has informed what aspects of more generic notions of life-quality, particularly with respect to mental and physical well-being are of particular relevance to exercise referral populations.

In addition, some of the items contained within the new scale are exclusive to exercise related life-quality. For example, items relating to the enjoyment of PA, confidence using gym equipment and exercising in a leisure centre with minimum support. In this respect the research outputs from the current study support what has been deemed as relevant measurement indicators for QoL by previous scale construction researchers (e.g., Skevington et al., 2004) but also extends this knowledge by providing insights into what additional indicators reflect those particularly sensitive to exercise-related life-quality. Arguably, these findings also support the rationale for the scales development as the items and factor structure of the final scale would suggest that a generic measure of life-quality would prove limited within an exercise referral environment. For example, a generic scale would be incapable of detecting aspects of injury pain and illness and PA facilitators that the current research has revealed as important to the perception of exercise-related life-quality.

The use of a Scree plot (Figure 13) that visually represented the Eigenvalues proved an effective approach to establishing the statistically robust final 22-item three-factor solution. Although it is recognised that Scree plots are reliant upon subjective interpretation, this visual representation of Eigenvalues helped to support the interpretation of a three-factor solution generated from the PCA. In this respect the analysis methods adopted by the current study support Cattell's (1966) assertion that Scree plots can be helpful in the process of factor reduction. Further analysis that forced the three-factor solution into a single factor indicated that all factors loaded significantly (>0.4) onto this factor and therefore has resulted in the added benefit of the final scale being able to be scored either as a single score, or as an expression of the three separate domains (sub-scale scores). This psychometric property of the final scale is further supported statistically by the results obtained for the convergent

validity of the new scale. Correlations between both the total scores and also between each of the domains of the new measure and the domains of the WHOQOL-BREF and the SEES were observed as was predicted based upon the item content of the new scale and the purpose for which the measure has been designed (i.e., the measurement of exercise-related life-quality). Employing statistical analyses in this way to support a scales ability to be scored either singularly or dimensionally have been utilised previously (e.g., Gianaros, Muth, Mordkoff, Levine, & Stern, 2001).

It is likely that in practice, exercise referral personnel with prefer to assess an overall rating of perceived exercise-related life-quality using a single score as this is usually more easily communicated to clients and medical professionals. It is also expected that this level of data is adequate for the scale's intended purpose for use within exercise referral settings. However, there may be occasions when assessing the three distinct domains of exercise-related life-quality will be beneficial. Perhaps if interviews with clients reveal that a particular area of the client's life poses current challenges to their exercise behaviour and is representative of one of the sub scale domains. For example, the domain injury pain and illness may be able to detect reduced scores for client's who are experiencing more pain than usual at the time of referral. Scores in this domain therefore, would require closer examination and future filed testing of the new scale would also allow for any changes in these reports to be detected over time (responsiveness). Although it is acknowledged that this is currently one of the limitations of the new scale. The flexibility of generating total and/or domain scores may be useful if using the new scale for research purposes because an assessment of both these scores would be beneficial to the research aims. Greater volumes of data generated in this way will also contribute to the future development of the scale, inform the specificity and sensitivity of the scales measurement also support the development of the conceptual framework that was used to construct it. This is addressed in more detail in Chapter VII.

Nine items that failed to meet the statistical criteria of item analyses were included in the PCA on the grounds of content validity (Streiner & Norman, 2008). Of these, five were included in the final scale. These items were Q10 regarding the enjoyment of PA, Q34 regarding perceived outlook on life, Q38 regarding perceived sense of well-being, Q47 regarding perceived life expectancy and Q49 regarding perceived overall life-quality all of which load onto the factor titled mental and physical well-being save for Q10 regarding the enjoyment of PA which loaded onto physical activity

facilitators. The *a priori* criteria set for the retention of items that failed item analysis was used to guide the process (i.e., those items that failed less criteria than others within the same conceptual framework category or were retained) (Chapter IV). The inclusion of these items in the final scale highlights the importance of the role of construct validity in contributing to the complete process of scale construction. Had these items have been rejected at the item analytical stage then these areas of exercise-related life-quality assessment would not be addressed by the final measure and arguably could have led to less breadth of assessment particularly with respect to the mental and physical well-being domain.

The resultant 22-item scale includes four reverse-scored items. Each of these items comprise the injury, pain and illness domain and high scores reflect aspects of exercise-related life-quality that reduce the potential for the QoL enhancing impact of regular exercise and PA for clinical populations (e.g., Milne, Guilfoyle, Gordon, Wallman, & Courneya, 2007; Motl & Snook, 2008). The general consensus is that the inclusion of reverse scored items encourages the respondent to pay careful attention to what is being asked before they provide an answer. A key figure within this area of research, Nunnally (1978) was a strong advocate for employing mixtures of regular and reverse-scored items within instruments. However, the literature demonstrates some debate regarding the use of reverse-scored items.

For example, Rodebaugh, Woods and Heimberg (2007) concluded that the reverse-scored items contained within the Social Interaction Anxiety Scale (SIAS; Mattick & Clarke, 1998) hinder the scale's psychometric performance and that these items are better considered as a measure of extroversion. From a population perspective, Conrad, Wright, McKnight, McFall, Fontana and Rosenheck (2004) concluded that the reverse-scored items included in the Mississippi Scale for Posttraumatic Stress Disorder (M-PTSD; Keane, Caddell, & Taylor, 1988) caused confusion for particularly 'disturbed' respondents and that their removal improved validity without loss of reliability. Nevertheless, the inclusion of reverse-scored items into the development of new scales is commonplace amongst psychometric researchers and factor analytical approaches tend to statistically group these items together as with the current study (e.g., Conrad, et al., 2004). Cognitive pre-testing (Chapter III) revealed that respondents had no difficulty in understanding these reverse-scored items. The inclusion of four reverse-scored items in the current scale allows for the

measurement of those lifestyle factors that impede upon enhanced perceptions of exercise-related life-quality specifically (i.e., injury, pain and illness) and should be considered a strength of the overall measure.

Support for the convergent validity of the final scale was provided from the correlations observed between the dimensions of both the WHOQOL-BREF and the SEES. As expected, the strongest positive correlations between the total scores of the new measure and the domains of the WHOQOL-BREF were physical health (r = .64; p <.01) and psychological health (r = .62; p <.01). The strongest correlation was observed between the test-measure total score and those from the positive well-being dimension of the SEES (r = .55; p <.01). These data would suggest that the total test score format of the new scale is sensitive to those aspects of exercise-related life-quality that are attributed to psychological and physical health. This is not surprising perhaps given that the mental and physical well-being domain of the new scale contains the most amount of items (n = 11) and that the new scale's primary purpose is to be considered QoL related measure.

An examination of the convergent validity of the new scale at sub-scale level indicated that the strongest positive correlations were between the mental and physical well-being dimension and the psychological well-being dimension of the SEES (r = .61; p .01), and the physical (r = .65; p <.01), psychological (r =.73; p <0.01), social (r = .56; p <.01) and environmental r = .64; p <.01) dimensions of the WHOQOL-BREF. The strongest negative correlation with this domain was observed for the personal distress domain of the SEES (r = .52; p <.01). This would suggest that the mental and physical well-being domain is particularly sensitive to the psychological and physical aspects of well-being but it is also interesting to note the strong correlations between the mental and physical well-being domain of the newly constructed scale and those of social and environment of the WHOQOL-BREF. As mentioned previously, item content of the mental and physical well-being dimension of the new scale is most closely related to generic measures of life-quality so it is not surprising perhaps that a generic scale correlates so well. Further field testing of the 22-item version of the scale will provide additional insights into how the mental and physical well-being dimension is related to more generic understandings of life-quality and may even help to inform the future development of the conceptual framework of exercise-related life-quality.

Sub-scale correlations between the injury pain and illness dimension of the new scale were strongly negatively correlated with the physical dimension of the WHOQOL-BREF ($r = -.71$; $p < .01$) which would suggest that the injury pain and illness dimension measures physical aspects of exercise-related life-quality. A strong positive correlation was observed between the PA facilitators dimension of the new scale and the physical domain of the WHOQOL-BREF ($r = .52$; $p < .01$). Moderate correlations were also observed between this sub-scale and the psychological domain of the WHOQOL-BREF ($r = .47$; $p < .01$) and the psychological well-being domain of the SEES ($r = .42$; $p < .01$). This would suggest that the PA facilitators' domain primarily measures facilitators related to physical ability. For example, there is an item contained within this domain that asks for a rating of confidence in physical ability. However, the moderate correlations with psychological domains of the WHOQOL-BREF and SEES would suggest that this domain also measures psychological facilitators of PA. Items relating to the perceived enjoyment of PA and confidence levels to exercise with minimum support would indicate that this is feasible.

The overall acceptability of the new measure was good. The frequency of missing responses did not exceed the level of 4% suggested by Fayers and Machin (2000) as an acceptable level for QoL assessments. The most frequent missing answer for the new scale was to the item that asks about levels of confidence using gym equipment (3%). It is likely that because respondents were recruited from a range of activities that this question was not always relevant to those who were participating in activities outside of a gym/fitness suite. Comparisons with the frequency of missing data for the WHOQOL-BREF (15%) and the SEES (4%) would suggest that the new measure is acceptable to exercise referral participants. Future field testing of the 22-item version could include recording the activities respondents were undertaking which may help to confirm if the most frequent missing data for the new scale was because of the range of activities undertaken by participants. A more complete review of the acceptability of the new scale is presented in the next chapter (Chapter VI).

The process of allocating the titles given to each of the three factors was guided by the recommendations of Pett, Lackey and Sullivan (2003) and a series of peer-debriefing sessions (Spall & Stephen, 1998) facilitated the process. The literature

offers little in terms of guidance for scale developers regarding this inductive process. However, considering the underlying meaning of the highest item loadings for each factor (Pett et al., 2003) coupled with a mindfulness of for whom the measure is intended and peer de-briefing sessions between the research team (Spall & Stephen, 1998) proved a simple yet effective method. As mentioned earlier, the final domain titles for each of the three factors were: mental and physical well-being, injury pain and illness and physical activity facilitators. The items contained within these domains have direct relevance for providing a greater understanding as to what constitutes exercise-related life quality and how it may differ from more generic measures if at all. Initial insights have been explored in the current chapter, more particularly with respect to the convergent validity data generated for the current study. These considerations are discussed in greater detail in the Discussion and Conclusions Chapter (Chapter VII).

5.5.1 Limitations of the Research

It is recognised that EFA is limited to providing an initial factor solution and that CFA is often performed to further assess the factor structure of newly developed scales. This is future research that is planned for the current scale under development (see Chapter VII). It is also recognised that because the same sample was used in the current phase of research as was used for item analyses, that the same demographic limitations reported earlier also apply to the current study (see section 4.5.1). Again, planned future research (Chapter VII) will provide opportunities to address these limitations.

5.6 Conclusions

A combination of item analytical and PCA statistical procedures coupled with content validity considerations, a mindfulness of the purpose of the final scale and for whom it was intended gave rise to a 22-item three-factor solution that has been shown to be reliable valid and acceptable. The new measure has retained a respectable breadth of assessment addressing areas of exercise-related life-quality affiliated to perceptions of mental and physical well-being, injury pain and illness and physical activity facilitators. These dimensions may provide valuable insights

into what constitutes exercise-related life-quality, how it may be measured and how it differs from more generic conceptions of life-quality. These considerations are addressed in Chapter VII.

The research reported in the current chapter completed the psychometric phases of test construction and revealed the final properties of the scale. These are reported in the following chapter (Chapter VI).

CHAPTER VI Properties of the Final Scale and Scoring Protocol

6.1 Introduction

The development of the current scale at item level involved item analyses that adopted a CTT approach to assessing the performance of and reducing the item pool. Initial assessments of reliability were also taken into consideration at this stage by an assessment of test-retest reliability and internal consistency. From an initial test pool of 50 items, 29 were carried forward into EFA. This procedure served to reduce the item pool further, determine the factor structure of the remaining 22 items and thus identify those components of life-quality that are particularly relevant for exercise referral. Test-retest reliability and validity analyses of internal consistency and construct validity at this scale level determined that the final scale was both reliable and valid in this respect.

The following sections of the thesis address the final properties of the scale, and how the measure should be scored and administered.

6.2 Appropriateness and Feasibility

The final scale takes approximately 10-15 minutes to administer. Cognitive interviewing revealed that respondents had no difficulty completing the measure under self-complete, interview or telephone interview conditions and therefore may be administered in any of these three ways. Consequently this increases the flexibility of the measures administration and increases the likelihood that exercise referral practitioners will be able to integrate the measure into routine practice. Some schemes have a scheme co-ordinator whose role it is to accept the referral from the health professional and contact the patient before referring on to a leisure centre, while some referrals are made straight to the leisure centre and contact with the patient is made directly by the exercise referral instructor. In instances where patients are contacted initially by a scheme co-ordinator, the measure will be particularly useful to administer over the telephone and in instances where the patient is contacted by the instructor, there is scope to introduce the measure either over the telephone or when the patient arrives for their first appointment.

In practice, patients referred with heart conditions or who have undergone heart surgery such as angioplasty or bypass surgery, for example will be signposted into specialised exercise classes referred to as phase IV cardiac rehabilitation. The training that exercise instructors receive to support patients within these exercise environments is much more specialised (e.g., British Association of Cardiac Rehabilitation; BACR) than exercise referral (e.g., the WRIGHT Foundation) and these patients are often considered as higher risk that those referred into exercise referral. Some schemes allow those who complete a programme of phase IV cardiac rehabilitation to move into exercise referral classes as an 'exit-route' option for those referred which is why it is becoming increasingly more common that patients with heart conditions can be found in exercise referral schemes. However, in the main, exercise referral and phase IV are considered as separate services. The scale was tested within both environments and also with heart patients attending exercise referral therefore, there is scope for exercise professionals to use the measure within exercise referral *and* phase IV cardiac rehabilitation settings which adds to the flexibility and feasibility of the measure's use.

The measure is relatively short, easy to use and score (see 6.3 for scoring protocol) and therefore, there is no training required to understand how to administer or score the measure. The primary investigator has good links with exercise referral schemes across the UK, particularly those that were involved in the development of the scale and in this respect, schemes will have access to support via contact with the primary investigator. The WHOQOL-BREF encourages those that use the measure either in research or practice to liaise with the developers at Bath University to help inform performance and future development. The same system will be adopted to support the future development of the current scale.

The current scale has been developed with patients and practitioners in mind at every stage of the research programme in an attempt to provide a measure that is relevant for practice. That said there is no reason why the measure could not be implemented within research settings as an alternative to the employment of generic measures of life-quality in studies that target exercise referral populations. The results generated from such research could also help with the future development of the measure.

6.2.1 Acceptability

The frequency of missing responses did not exceed the level of 4% suggested by Fayers and Machin (2000) as an acceptable level for QoL assessments. Data were missing most frequently in response to levels of perceived confidence regarding the use of gym equipment (n = 8). It is likely that this item was missed by some respondents because not all those who completed the test measure were referred to a leisure centre that required the use of gym equipment. Some respondents were referred into group exercise classes or swimming for example. Items that had the lowest level of missing data included those regarding outlook on life (n = 1), self-esteem (n = 1) and sleep quality (n = 1). The item regarding overall QoL was responded to by all 278 participants (100%).

In addition to an analysis of the frequency of missing data, support for the acceptability of the test measure can be drawn from the cognitive pre-testing data. Respondents found completing the measure to be acceptable either under self-complete, interview or telephone administration conditions. The test version of the measure included 50 items. Therefore, it is expected that the shortened 22-item final version of the scale will pose even less burden to respondents than the 50-item version tested.

6.2.2 Precision

Before taking any scoring protocol into consideration, the current scale reflects a raw range of scores from 22-110 with 110 representing the highest exercise-related QoL score possible. In the sample of respondents used for the psychometric phases of research (N = 278), the mean total scores were (77.04, SD ± 11.46) and total scores ranged from 45 to 105. This data indicates a good level of variability, there is no bunching of scores at either the upper or lower level of scores and therefore the scale is not limited by ceiling or floor effects.

The response options are presented as a 5-point Likert scale and designed from previous research that has been able to demonstrate the suitability and precision of such response scales for health assessment (Skinner & Tucker, 1999). The final scale

represents a measure of adequate length so as not to be burdensome upon patient or practitioner, yet is broad enough to cover those aspects of life-quality that are relevant to the population for whom the measure is intended. As discussed previously, the 22 items represent a robust three factor structure which was informed by a conceptual understanding of exercise-related life-quality which was previously lacking. This relates to the precision of the final scale because these statistically robust and representative three factors underpin the measure's ability to be scored and communicated either as an expression of these dimensions or as single global score. Of course this also further contributes to the appropriateness and feasibility of the measure for exercise referral and public health practitioners because a single global score of exercise-related life-quality is more easily communicated between professionals and patients than a number of dimension sub-scores.

6.2.3 Reliability

The internal consistency and test-retest reliability were calculated during the final phases of scale construction (Chapter V). The internal consistency of all 22-items included in the three-factor solution was good with a Cronbach's alpha of 0.88 and alpha could not be improved by deleting any of the 22 items included in the final scale. The internal consistency (Cronbach's alpha) of the mental and physical well-being dimension was 0.87, injury pain and illness was 0.81 and the physical activity facilitators was 0.77. Additionally, all 22-items met the item-total correlation criteria of >0.3 (Skevington et al., 2004). Test re-test (intraclass correlation) at scale level indicated a significant level of reliability ($r = .72$; $p < .01$).

An examination of the convergent validity of the new scale at sub-scale level indicated that the strongest positive correlations were between the mental and physical well-being dimension and the psychological well-being dimension of the SEES ($r = .61$; $p < .01$), and the physical ($r = .65$; $p < .01$), psychological ($r = .73$; $p < .01$), social ($r = .54$; $p = 0.01$) and environmental $r = .64$; $p = 0.01$) dimensions of the WHOQOL-BREF. The strongest negative correlation with this domain was observed for the personal distress domain of the SEES ($r = .52$; $p < .01$).

6.2.4 Face and content validity

Cognitive pre-testing revealed that the measure has good face and content validity. Any items that were misinterpreted or required clarification were modified following this phase of scale construction research and although these modifications were minimum (six questions were re-worded, one question was deleted and one was split into two separate questions), cognitive pre-testing enhanced the face validity of the measure in this respect by eliminating problem items early on. Additionally, cognitive pre-testing was able to demonstrate that the overall presentation and formatting of the measure was appropriate and did not cause any difficulties regarding respondent's ability to complete or understand the measure.

Content validity was considered throughout the initial scale construction and psychometric phases of research. The rich level of focus group data were able to generate a wide range of items that were representative of the dimensions of exercise-related life-quality reported by the focus group participants. Coupled with demographically representative samples of participants at each stage of the research programme, this enhanced content validity in this respect. Specifically, because the ill-health conditions with which participants were referred were wide ranging yet typical of those for which exercise is prescribed. In addition, the age range of participants was also representative of the population for whom the final scale is intended.

Content validity was also demonstrated at a further level. A conceptual framework of life-quality was developed from the focus group data which informed item construction and also the iterative decision making process of item retention and rejection during item analyses. The factor structure of the final 22-items is representative of all the components that comprise the conceptual framework. This also contributes to the content validity of the scale as this contextual framework is a reflection of the reports made by focus group participants regarding exercise-related QoL. A more detailed critique of the content validity of the final scale in terms of the factor structure, conceptual framework and how these provide insights into what is actually measured by the final scale can be found in Chapter V.

6.2.5 Construct Validity

Chapter V reports on the convergent and known groups validity components of construct validity. As expected there were strong correlations observed between the test-measure total score and the dimensions of physical health ($r = .64$; $p < .01$) and psychological health ($r = .62$; $p < .01$) of the WHOQOL-BREF. Moderate correlations were observed between the test-measure total score and social relationships ($r = .47$; $p < .01$) and environment ($r = .56$; $p < .01$). These data supported the *a priori* predictions regarding the strength and direction of correlations between the test measure and the domains of the WHOQOL-BREF and the SEES and thus support the convergent validity of the test measure.

Additional support for the convergent validity of the test measure was offered from the observed strong correlation with the test measure total score and those from the positive well-being dimension of the SEES ($r = .55$; $p < .01$). This was further supported by observing moderate negative correlations with the psychological distress dimension of the SEES ($r = -.45$; $p = 0.01$). Sub-scale correlations between the injury pain and illness dimension of the new scale were strongly negatively correlated with the physical dimension of the WHOQOL-BREF ($r = -.71$; $p < .01$) which would suggest that the injury pain and illness dimension measures physical aspects of exercise-related life-quality. A strong positive correlation was observed between the PA facilitators dimension of the new scale and the physical domain of the WHOQOL-BREF ($r = .52$; $p < .01$). Moderate correlations were also observed between this sub-scale and the psychological domain of the WHOQOL-BREF ($r = .47$; $p < .01$) and the psychological well-being domain of the SEES ($r = .42$; $p < .01$). These data also supported the *a priori* predictions regarding the strength and direction of correlations between the sub-scales of the test measure and those of the WHOQOL-BREF and the SEES and therefore also support the convergent validity of the test measure. Chapter V documents the convergent analyses in more detail including the correlation predictions and a critique of how correlations between the test measure (both at total score and domain level) may inform what indices of exercise-related life-quality are being measured within each domain of the new scale.

During the final phases of psychometrics, the known group component of construct validity was also calculated (Chapter V). The current test measure has been designed to be sensitive to the effects of exercise for clinical populations. As such significant

differences in test scores for those who have just started their programme of exercise and those who have completed would be expected. The largest significant difference in mean total scores was between those that had just started their exercise referral programme and those that had completed ($r = .01$; $p < .05$). There was also a significant difference in mean total scores for those that had just started their referral programme and those that were midway through ($r = .03$; $p < .05$). Differences in mean total scores for those respondents midway through their referral programme and those that had completed were also observed, although these differences were not significant ($r = .55$; $p < .05$). Although further field testing would need to be undertaken to assess the scales ability to detect change over time using the same groups of respondents (responsiveness), this provisional data suggests that the current test measure is concentrating on those areas of life-quality that are particularly sensitive to exercise. To assess this accurately, a longitudinal study that measured the same respondent's exercise-related life-quality at programme start, midway through and upon completion would have to be undertaken and as such offers a suggestion for future research.

6.2.6 Reproducibility

Respondents (N = 278) were asked to complete the test measure on two separate occasions, one week apart. Chapter IV details the specific methods employed and the rationale for the time lapse between each test measure completion. The data between time one and time two was then assessed at both item and scale level. The test-retest reliability of items assessed using weighted kappa indicated that all 50 test items met the criteria set by Landis and Koch, (1977). Similarly, the test-retest reliability of the final 22 items at scale level as assessed by intraclass correlation indicated a significant level of reliability ($r = .72$; $p < .01$).

6.2.7 Interpretability

The current measure is a new measure of exercise-related life-quality and as such the availability of data to assess the interpretability of the measure is limited. Section 6.4 addresses how to score the new measure. The data presented in section 6.4 regarding the percentile scores of the respondents included in the current psychometric phases of research (N = 278) may provide some insights into the interpretability of the

measure. However, further research needs to be undertaken to generate comparative data from exercise referral schemes across the UK. As mentioned previously, it would be particularly valuable to gather same-group longitudinal data to assist with furthering understanding regarding the responsiveness and interpretability of the new measure.

6.3 Appropriate Scale Title: The ER-QLS

The literature has little to offer in terms of guidelines for the appropriate naming of new scales. Significantly more attention is paid to the critical stages of item development and psychometrics. Although a review of the titles given to health and QoL assessment scales seems to indicate that the full title of the measure is often abbreviated and it is this abbreviation to which the scale is referred. For example, the Cancer Rehabilitation Evaluation System Short Form (CARES-SF) (Schag et al., 1991), the Centre for Epidemiologic Studies-Depression Scale (CES-D) (Radloff, 1977) and the Short Form 36 (SF-36) (Brazier et al., 1992) to name but a few. Some QoL measures include a reference to QoL in the title. For example, the WHOQOL-BREF (Skevington, et al., 2004) and the Eating Disorders Quality of Life Scale (EDQLS) (Adair et al., 2010).

What is perhaps of greater importance is that the title of a new scale reflects that which it purports to measure. The difficulty regarding the definition (e.g., Cohen, Mount, & MacDonald, 1996) and the measurement of life-quality (e.g., Allison, Locker, & Feine, 1997; Holmes, 2005) has already been addressed (Chapter I) and will be addressed within the context of the new scale in greater detail in the General Discussion (Chapter VII). The items included in the current scale were specifically designed to address either causal or effect indicators of exercise-related life-quality which has been identified as a particularly important consideration when developing a QoL scale (Fayers et al., 1997). Any items that were not considered causal or effect items were rejected during the early phases of item construction. The new scale has also demonstrated convergent validity with a generic measure of life-quality (WHOQOL-BREF; Skevington et al., 2004) which would suggest that the new scale measures what is purported to be QoL by an existing scale but that is particularly sensitive to exercise referral as demonstrated by known group validity.

Therefore, the title Exercise Referral Quality of Life Scale offers a clear description of what the scale is intended to measure and an abbreviation to ER-QLS is concurrent with usual scale development practice. Therefore, from this point forward the newly constructed and validated measure will be referred to as the ER-QLS.

6.4 Scoring the ER-QLS

Factor analysis of the final 22-items revealed an underlying three-factor solution that may be scored by each dimension or as a single global score (Chapter V). It is usual when scoring the responses to test items to simply sum the total value of responses. However, this method assumes that all the items are equally important in their contribution to the test (Rust & Golombok, 2009). One method of addressing this is to 'weight' the importance of each item so that those items that are considered of most importance in the scale are given a higher score weighting. Rust and Golombok (2009) have noted that this is rarely done and that a standard total score equal weighting approach has many practical advantages including increasing the ease of evaluating tests, providing instructions on scoring and setting standards. While Streiner and Norman (2008) argue that weighting of items is "rarely worth the trouble" (Streiner & Norman, 2008 p. 163).

A further consideration when developing a scoring protocol for a newly validated measure is whether comparisons to the scores from other, related measures will be required in practice. Adding up the sum of items to derive a total score may prove problematic if the comparison scale scores are reported on a different metric (Streiner & Norman, 2008). Although the authors argue that this is not so much of a difficulty if the scale is targeting a brand new area. Given that the ER-QLS has been designed to offer a quick and easy way of assessing exercise-related life-quality within an environment where reducing both practitioner and respondent burden are important, the ease of scoring the instrument is equally as important as its development. It is also important to note once more that exercise referral personnel are not typically trained researchers, nor are they required to collect and analyse complex data. If this level of data is required, it is more usual that a scheme co-ordinator or an academic institution will undertake this work.

To that end, to facilitate the ease of administration and the likelihood that the ER-QLS will have a key role in the evaluation of schemes the scoring protocol was designed to generate either a single total score or three sub-scale scores that can be calculated either by hand or with the use of statistical software packages such as SPSS. This will be particularly beneficial for the instances where larger schemes require the analysis of greater volumes of data. An SPSS syntax file for scoring can be found in Appendix 20 and will be made available to practitioners. The weighting of items either through *a priori* hypotheses or with the use of importance scales was not included in the scoring protocol because of the advantages of omitting these highlighted by Rust and Golombok (2009) and Streiner and Norman (2008). Furthermore, items were not weighted so as to avoid the added paperwork generated from importance scales, to limit respondent burden and administration time. As Streiner and Norman, (2008) have identified, the cross-comparison of scores between the ER-QLS and the data generated from similar measures is of little importance as there is no other QoL measure that assesses the exercise-related life-quality of exercise referral participants and so integrating a conversion method to allow for cross-score-comparisons was not necessary.

The scoring of the ER-QLS follows the guidelines provided for one of the construct measures utilised in the scales development and validation. The WHOQOL-BREF (Skevington, et al., 2004) makes use of the same response options as the current measure (Skevington & Tucker, 1999) and provides clear guidelines as to how the measure should be scored. The current measure adopts a scoring system of 1-5 for each item with a higher item and total score denoting a better perception of exercise-related life-quality. However, the scale also includes four reverse-scored items that need to be factored into the calculation. Drawing from the WHOQOL-BREF guidelines, the ER-QLS should be scored in the following manner:

For a single total score:

a) Check that all the scores from the 22 items have scores ranging from 1-5.
b) Reverse the four negatively framed items (Q12, Q13, Q14 Q15) so that 1 = 5, 2 = 4, 3 = 3, 4 = 2 and 5 = 1.
c) Compute the total score by adding all the response totals.
d) Total scores are presented in the range of 22-110. The higher the score, the greater the perceived exercise-related life-quality.

To score each of the three dimensions:

a) Check that all the scores from the 22 items have scores ranging from 1-5.
b) Reverse the four negatively framed items (Q12, Q13, Q14 Q15) so that 1 = 5, 2 = 4, 3 = 3, 4 = 2 and 5 = 1.
c) Compute the dimension mental and physical well-being by adding Q1+Q2+Q3+Q4+Q5+Q6+Q7+Q8+Q9+Q10+Q11
d) Compute the dimension injury pain and illness by adding
Q12+Q13+Q14+Q15
e) Compute the dimension physical activity facilitators by adding
Q16+Q17+Q18+Q19+Q20+Q21+Q22
f) Total scores for the dimension mental and physical well-being are presented in the range of 11-55. Total scores for the dimension injury pain and illness are presented in the range of 4-20. Total scores for the dimension physical activity facilitators are presented in the range of 7-35. In each case the higher the score, the greater the perceived exercise-related life-quality in each domain.

A critical consideration for new scale development is how to attribute meaning to the resultant scores. A score of 75 out of 100 on a QoL scale tells us nothing about the person's QoL. All we know is that the person's QoL is better than someone who has scored 40 and not as good as someone who has scored 85. Arguably the most efficient way of achieving robust interpretations of scores would be to conduct a longitudinal study with large sample sizes in order to establish group or population norms and to express scores as percentiles of these norms. It is commonplace for scale scores to be expressed as percentiles, particularly in medicine whereby height and weight charts are interpreted upon this basis (Streiner & Norman, 2008). Such a large scale study to determine population based norms is beyond the scope of the current programme of research and so based upon the guidance provided by Streiner and Norman (2008) meaning to the scores generated from ER-QLS or thresholds has been calculated by determining the 25^{th}, 50^{th} and 75^{th} percentiles of the cohort used for the current psychometric studies (N = 278; M = 76.00; SD ± 16.41). Table 34 details these values.

Table 34. Percentile scores for the ER-QLS

25^{th} Percentile	50^{th} Percentile	75^{th} Percentile
62.50	76.00	89.50

6.5 Instructions for Administrators

The following guidelines have been developed from the recommendations of Streiner and Norman (2008) and Rust and Golombok (2009). The experiences of administering the ER-QLS, perhaps more particularly during the cognitive pre-testing phase of research (Chapter III), informal feedback from the exercise referral schemes involved in the later phases of research and observations made by the primary investigator have also informed the following administration instructions:

a) Cognitive interviewing revealed that there were no differences in respondents ability to complete the scale across three methods of administration: self-complete, interview or telephone administration. Therefore, the ER-QLS may be self administered or interview-administered in person or over the telephone. The standardised instructions provided on the first page of the scale should be read aloud to respondents in instances where the instrument is interview-administered either in person or over the telephone.

b) If the measure is to be administered in person, ask the respondent if they would prefer to self-complete or have the questions and response options read aloud to them. If during administration the respondent wishes to swap their method of completion allow them to do so.

c) In instances where the measure is to be self-completed, allowing the respondent to complete without a professional present at the time should be avoided. Respondents often elaborate on the responses that are given, providing opportunities to explore their relationship with physical activity and exercise and specific barriers and facilitators to participation, for example. Having an exercise referral professional present during administration also allows any queries respondents may have or

difficulties they may encounter to be addressed and facilitates the development of good rapport between the client and exercise referral professional.

d) If a respondent fails to understand a question, in the instance of self-completion ask them to read the question and response options again or read these aloud for them. If a respondent fails to understand a question and/or response option while the measure is being interview-administered (in person or over the telephone) repeat the question and the response options again. Do not change the wording of any question or corresponding response options. If a respondent is unable to answer a question, move on to the next one.

e) If a respondent wishes to provide an answer that falls between two points on the response scale (e.g., between 'not at all' and 'not much') encourage the respondent to choose the response that is closest to their desired answer. Assure the respondent that there is no 'correct' answer and they should attempt to describe how they feel.

f) If using SPSS to manage respondent data and to calculate total scores, a SPSS spreadsheet and syntax file can be made available to you by contacting the primary investigator (CH). The syntax file automatically checks, recodes data and computes total QoL scores and will allow you to communicate large volumes of data to stakeholders and funding bodies (Appendix 20).

6.6 Discussion

This chapter has presented the final properties of the ER-QLS, determined an appropriate scoring protocol and offered instructions for administrators. The overall aim of the complete phases of research was to assist with the current shortfall in the psychological evaluation of exercise referral schemes and to produce a measure of life-quality that is particularly sensitive to exercise referral. The data generated from cognitive pre-testing and the psychometric properties of the scale suggest that this aim has been met.

The test version of the final scale was well received by respondents during both the cognitive pre-testing and the psychometric phases of research. Although it is

acknowledged that the final 22-item version of the ER-QLS will need to be tested to inform its performance and acceptability, there are lessons to be learned from the outputs of the current research that provide insights into the rigor and appropriateness of the final scale properties. The final scale was devised from distinct and purposeful phases of research that specifically addressed a) the generation of items, b) construction of items, c) testing of items and d) psychometric performance at scale level. As such the final scale properties are best judged in light of the procedures responsible for its construction. These are critiqued below.

6.6.1 Item Generation

The end point of a new scale has to be that it is representative and relevant for the population for whom it is intended. Therefore, the source of items is of critical importance. No amount of statistical manipulation can account for poorly chosen questions and patients rather than clinicians (or what is often termed as the 'expert opinion' approach to item generation) have been deemed as "an excellent source of items." (Streiner & Norman, 2008 p. 18). Items for the ER-QLS were generated from the reports of exercise referral participants who had completed their initial 12 weeks of exercise and were therefore in a favourable position to communicate their experience of exercise with specific reference to perceptions of life-quality.

The source of items is important not only with respect to who will generate the data but also how. One of the strengths of the focus group method has been identified as the flexibility to use focus groups as a stand-alone method or as employed by the current thesis, in combination with quantitative techniques as part of a multi-method project (Wilkinson, 2003). Furthermore, the analysis of focus group data is not bound to any particular theoretical frame work and the primary purpose of analysis is to elicit the ideas, understandings and opinions of individuals (Wilkinson, 2003). A focus group approach allows for participants to reflect upon their own interpretations and experiences based upon the reports made by others in the group (Kitzinger, 1994, 1995), and in turn provides a richness or depth of data to interpret. In this respect, focus groups are much more suitable then individual interviews for the purposes of item generation and played a key role in this initial phase of research. There is no doubt that the purposeful selection of appropriate participants and the focus group method employed contributed to the appropriateness, volume and depth of data generated which was particularly favourable for item construction.

6.6.2 Item Construction

Constructing qualitative data into meaningful questions is no arbitrary task. The construction of items for the ER-QLS was dependent upon the quality and appropriateness of the data that informed the process. The construction of poorly worded or ambiguous questions during this phase of research would have impacted negatively upon the subsequent psychometric phases of scale development and resulted in a scale that was unsuitable for use. Therefore, a considerable amount of time was spent translating the focus group data into items that were likely to perform well and this phase of research was guided by a number of key texts (e.g., Brace, 2004; Foddy, 1993; Hauge, 1993; Oppenheim, 1992). Iterative phases of peer de-briefing (Spall & Stephen, 1998) were undertaken whereby the test item pool was reviewed by the research team and suggestions for item re-wording or question order for example, were made until it was agreed that the item pool reflected a rigorous set of questions that were suitable for cognitive pre-testing.

6.6.3 Cognitive Pre-testing

The items that comprise a new scale under construction are not always pre-tested. Perhaps this is due to an over-reliance on the ability of psychometric procedures to detect those items that do not perform well enough to be retained. Uncertainty regarding the procedures to be undertaken may also contribute to the lack of consistency in utilising this phase of research (Presser, 2004). Cognitive pre-testing was considered as an integral and important part of the development of the ER-QLS because not only did such procedures identify those questions that required modification or deletion, but the process of undertaking pre-testing and the results generated provided valuable insights into face validity, levels of acceptability and how the final measure may be administered. This is beyond the scope of statistical approaches to item and scale performance and as such the results generated from the current study highlights the value of pre-testing as a consideration for other researchers.

Specifically, cognitive pre-testing revealed that six questions required re-wording, one question was deleted and one was split into two separate questions (Chapter III).

One of the questions that was amended during this phase of research is included in the final 22-item set. “How often do you feel you have to manage any symptoms of any illness you have?” Originally included “if any” at the end of the item and was subsequently removed following pre-testing. Problems with item wording would have carried forward into psychometric analyses had the question not have been improved during pre-testing. As a consequence this item may not have performed adequately and may have been eliminated from the final item set. In turn, this would have resulted in the frequency of illness symptom management not being measured by the ER-QLS or contributing to conceptual understandings of exercise-related life-quality.

In addition to ensuring the adequate performance of items and response scales in the manner described above, cognitive pre-testing had further value in contributing to the properties of the final scale. This phase of research allowed the face validity of the test version of the scale to be considered and also provided insights into appropriate methods of administration. The final scale is formatted in the same way as the test scale that was subject to pre-testing which demonstrated a good level of face validity (Chapter III). Therefore, it is feasible to assume that a fewer number of items (i.e., 22 rather than 50) will facilitate the acceptability and face validity of the ER-QLS in addition to reducing the level of respondent burden created by the 50-item test version. The three methods of administration (i.e., self-complete, interview and telephone) revealed no method-related difficulty with the interpretation of the test scale. Therefore, it is also feasible to suggest that the shortened 22-item final scale may be administered in each of these three ways. This is a particular strength of the properties of the ER-QLS because it enables the measure to be incorporated more easily into the patient pathway designs that are typical of exercise referral schemes within the UK. Had cognitive pre-testing not been undertaken or if the approach employed had limited testing to only a single method of administration such insights into how the ER-QLS should be administered would have been limited.

6.6.4 Item and Scale Analyses

A CTT approach to item analyses enabled the item pool to be reduced to a more appropriate number and in doing so identified those items that statistically performed the best (Rust & Golombok, 2009). The conceptual framework of exercise-related life quality that was devised as an output of the focus group data (Chapter II) served as a key instrument in ensuring that the ER-QLS adequately covered the domain under investigation. Such considerations have been recognised as contributing to the content validity of a scale under construction (Streiner & Norman, 2008). These content validity considerations operated alongside statistical analyses to ensure adequate representation of the contextual framework and reduce the likelihood of developing a scale that was too limited in the breadth of its measurement (Klein, 2000).

The psychometric procedures undertaken at scale level mean that the reliability and validity of the ER-QLS can be expressed in a number of ways and have been presented in the current chapter. The internal consistency of the total scale items and each of the respective domains as assessed by Cronbach's alpha was good. The test-retest reliability at scale level (intraclass correlation) indicated a significant level of reliability ($r = .72$; $p < .01$) and the construct validity of the ER-QLS was supported by predicted correlations with the WHOQOL-BREF and SEES (convergent validity) and also the known group component of construct validity.

Overall the ER-QLS displays sound psychometric properties and has been developed from purposeful phases of research, each designed to ensure the rigour of the final scale. The total number of reliability and validity analyses undertaken in the development of the ER-QLS represent two measures of reliability (internal consistency and test-retest) and three measures of validity (face, content and construct). It is recognised that further work will need to be undertaken to test the responsiveness of the measure in particular. However, the reliability and validity presented for the final properties of the ER-QLS is comparable to previous scale construction studies (e.g., Burckhardt & Anderson, 2003; Hilari, Byng, Lamping & Smith, 2003) and it is typical that further work is undertaken as new measures are introduced into the environment for which they were intended.

In assessing the final scale properties of the ER-QLS it is important to acknowledge the limitations. The data presented in the current chapter represents that which was generated from the initial 50-item test version of the ER-QLS and the final 22-item measure has yet to be field tested although plans are underway to carry out this research. The respondents involved in the testing of the ER-QLS were selected to be as representative as possible of the population for whom the measure is intended. This included, age, sex, ill health conditions and also took stage of referral into consideration. In this respect, respondents were considered as a homogenous group, representative of those who are typically referred into exercise programmes. It is acknowledged, however that the diversity in health conditions in particular may justify future research that investigates the performance of the ER-QLS based upon specific health conditions. As mentioned previously, there are QoL tools available for a number of disease states (e.g., Doward et al., 2003; Kolotkin, et al., 2001) and there may be scope to administer the ERQOLS alongside these more disease-specific measures to gain insights into the performance of the ER-QLS in this respect.

On a related point, the creation of group norms for the individuals who represent exercise referral populations (e.g., age, sex, health condition, ethnicity) would add to the interpretability of the ER-QLS and as already acknowledged, future work is needed regarding the responsiveness of the scale.

The ERQOL is unique in that it represents the first measure of life-quality to be designed specifically for exercise referral. While this should be considered a strength of the current research it also means that it is not possible to compare the performance of the ER-QLS with other measures designed for the same purpose. A critique of the similarities and differences between the ER-QLS and the generic measure of life-quality the WHOQOL-BREF was provided at item level (Chapter IV). The following chapter (Chapter VII) expands upon these initial observations and addresses the larger methodological issues involved in the development of the ER-QLS. In particular how the ER-QOL differs from more generic measures of life-quality, how the development of the ER-QLS may inform a greater understanding of the exercise life-quality relationship and the lessons learned regarding those reports made by focus group participants that were neither causal or effect indicators of life-quality and were therefore excluded from the measure.

6.7 Conclusions

The final scale properties of the 22-item ER-QLS generated from data during the testing phases of research suggest that the measure is reliable, valid and relevant to exercise referral. The final scale represents three distinct dimensions of exercise-related life-quality namely: mental and physical well-being, injury pain and illness and physical activity facilitators. The measure is easy to complete and pre-testing of the test measure suggested that it may be administered by self-complete, interview method or over the telephone. The ER-QLS is also easy to score and is able to offer a single global score of exercise-related life-quality or three separate scores for each dimension.

Coupled with the statistical approaches used to establish the validity and utility of the ER-QLS, cognitive pre-testing played a key role in informing the specific utility of the scale. A representative sample of participants were able to report directly regarding the suitability of the test scale for use within structured exercise settings and in the main, these reports were favourable. Arguably, if the people for whom the measure is intended cannot respond to the items contained within a scale then any number of statistical procedures will be unable to remedy this situation.

The level of critical attention given to the methods employed at each stage of the design, construction and testing phases of the ER-QLS is a particular strength of the scales final properties and the results would suggest that the primary aim of the thesis has been met.

CHAPTER VII Discussion and Conclusions

7.1 Introduction

The principal aim of this thesis was to develop and validate a QoL measure that would be particularly sensitive to, and relevant for exercise referral schemes operating within the UK. A series of studies that have utilised both qualitative and quantitative approaches to research have been undertaken to fulfil this principle aim. This chapter addresses some of the fundamental considerations of the research process and presents the key outcomes and contributions of the current thesis to scientific understanding. This chapter also discusses the relationship between these considerations and findings in light of existing empirical and theoretical work and in terms of their implications for policy and practice.

7.2 The ER-QLS and the Measurement of Life Quality

For practitioners QoL has come to describe an umbrella term for the measurement of what Greenley, Greenberg, & Brown (1997) have termed indicators of how a client is doing in a range of life situations. Given the relationship between PA, exercise and a variety of mental, physical, social and environmental health parameters (e.g., DH, 2004a), it makes sense that exercise referral practitioners would want to capture this and in a way that is easily communicated to others. This practice-based necessity was one of the key contributors to the rationale for the development of the ER-QLS and the psychometric properties of the scale would suggest that this aim has been met. Although despite this measurement necessity, it is important to acknowledge and re-visit the persistent lack of agreement about the meaning of QoL. Chapter I presented some of the key difficulties with QoL research namely: the definition (e.g., Cohen, et al., 1996) and the measurement of the concept (e.g., Allison, et al., 1997; Holmes, 2005). As Hunt (1997, p.205) has argued: "it is the first tenet of any scientific enterprise, where measurement is to be attempted, that the object of measurement be precisely and meaningfully defined and the measuring instrument be appropriate and valid for the task." The section to follow will address the challenges surrounding the measurement of QoL as is exemplified by Hunt's (1997) argument. Specifically, this section of the thesis will explore debate surrounding the measurement of life-quality in more general terms and also with respect to the implications for the ER-QLS.

One example of defining and therefore attempting to understand QoL has been provided by Lehman (1983) who described life-quality as the sense of well-being and satisfaction experienced by people under their current condition. The ER-QLS contains a specific question that measures a respondents perception of well-being. There are also a number of items that take the respondents current health, social and environmental conditions into consideration that are grouped into one of three dimensions: mental and physical well-being, injury pain and illness and PA facilitators. In this respect the ER-QLS has the capacity to measure QoL in the manner described by Lehman (1983).

The difficulty arises when we consider other definitions of life-quality that while may present similarities may also include key differences. For example, Hansson, (1999) has described QoL as a persons satisfaction with various *a priori* set of life domains compiled from an expert or hypothetical perspective. Does the 'bottom up' approach to the development of the ER-QLS that included patient reports in preference to the expert opinion approach to item generation mean therefore, that the ER-QLS no longer measures QoL as previously defined by Lehman (1983)? There are a number of other definitions that could be included to exemplify this point but the underlying message is that a lack of consistency or agreement in terms of how QoL should be defined or understood makes any measurement of it almost an impossible task. The differing content of generic measures of life-quality further highlights the transience of the concept. It could be argued that if life-quality could be universally defined and measured that there would be a single reliable and valid measure designed to capture it. However, this is not the case and in some respects reflects the very nature of scientific enquiry which is to offer differing hypotheses and critically examine them. In this respect the level of success attributed to a scale's ability to measure QoL can only be assessed by the criterion used to define life-quality by the scale in question. As a consequence, there have been calls to undertake research that attempts to define and understand QoL more fully in order to ascertain if it can be meaningfully measured (e.g., Hunt, 1997).

The development of health-related QoL measures go some way perhaps to tackling these difficulties as they attempt to define what aspects of global life-quality are relevant to the population for whom the measure is intended. In some respects this approach is designed to filter out those more generic aspects of life-quality so as to produce a scale that may address one of the further difficulties with QoL

measurement highlighted by Hunt (1997) which was clinician's lack of knowledge regarding what exactly is being evaluated. The ER-QLS is not intended to be relevant to a single ill-health condition and therefore there is a limited degree of ill-health or disease state specificity included within the items that comprise the scale. However, the notion of identifying what aspects of life-quality are being measured by a scale purporting to be a measure of life-quality has direct implications for the final properties and appropriate use of the ER-QLS. Therefore, the following section explores how the current research has informed what indicators may contribute to exercise-related life-quality from both an assessment and a conceptual perspective.

7.3 Exercise-Related Life-Quality

The ER-QLS is the first of its kind to be developed with exercise referral clients in mind and each phase of the development of the new scale was designed to target aspects of life-quality that are specifically related to exercise behaviour. Face and content validity considerations addressed through cognitive pre-testing and the construction of a conceptual framework of life-quality supported the acceptability and relevance of the measure in practice. This is of particular importance given that it has previously been reported that many clinicians feel frustrated and confused when asked to administer questionnaires which appear to have little or no rationale for being termed measures of life-quality (Hunt, 1997). This observation echoes what has been discussed in the previous section regarding the challenges of working with such a transient concept as QoL but the argument made by Hunt (1997) also demonstrates the importance of defining what a scale does and does not measure and how this does or does not relate to life-quality. Therefore, this section of the thesis will focus upon what the development of the ER-QLS can tell us about *exercise-related* life-quality in particular. Specifically, what the current research has revealed as being possible indicators of exercise-related life-quality with respect to a conceptual framework, and the factor groupings of the ER-QLS and what implications this has for measurement. In addition, a critical appraisal of what the ER-QLS does not measure, the implications for administration and for exercise referral as a whole is presented in the section to follow (7.4).

Angermeyer, Holzinger, Kilian, & Matschinger (2001) asserted that generic QoL assessments such as the WHOQOL-BREF cannot replace specific instruments which

cover the characteristic living conditions of a particular patient group. Alongside some of the issues surrounding the measurement of generic life-quality that have already been addressed, this viewpoint perhaps explains why there are so many measures of life-quality available that have been designed to focus upon a specific population or disease state (e.g., Kolotkin, et al., 2001). However, such measures have been accused of not providing clear criteria with respect to the specification of what is actually being measured and how the instruments can claim to measure it (Hunt, 1997). This has lead Hunt (1997) to propose that theory development should be encouraged perhaps even at the expense of questionnaire development. The distinct phases of scale construction undertaken for the current thesis have allowed for these very issues regarding what is actually being measured and also some theoretical development to be addressed. Further theory development could be encouraged by adopting the conceptual framework of exercise-related life-quality (Figure 8) as a theoretical underpinning for future studies that fall within this area of research interest. It is likely that such an approach to future research will help to alleviate some of the challenges regarding measurement uncertainty identified by Hunt (1997).

The analysis of data generated from five focus groups comprising a purposeful and representative sample of 23 participants was used to develop a conceptual framework of exercise-related life quality. An examination of this framework will provide insights into what aspects of life quality are particularly relevant within the context of exercise and exercise referral. Furthermore, reviewing this framework in light of what is measured by the ER-QLS may also help to bridge the gap between understanding what is intended to be measured and that which is actually measured. Chapter II presented the conceptual framework that identified leisure facilities and lifestyle PA, discomfort and fitness, mental and physical well-being and confidence and motivation as important indicators of life-quality that are related to exercise. Chapter V then detailed how this conceptual framework later mapped onto the resultant domains of the ER-QLS namely: mental and physical well-being, injury pain and illness and physical activity facilitators. In short, all of the components that comprise the conceptual framework are represented by the resultant three-factor structure. All the indicators that were conceptually grouped into mental and physical well-being were statistically grouped under the domain mental and physical well-being (11 items). The indicators that were conceptually grouped into leisure facilities and lifestyle PA were split between the injury pain and illness domain (1 item) and

physical activity facilitators domain (2 items). The indicators that were conceptually grouped into discomfort and fitness were split between the domains injury pain and illness (3 items) and physical activity facilitators (1 item). The indicators conceptually grouped into confidence and motivation were all statistically grouped into the domain physical activity facilitators (4 items).

The differences between conceptual and statistical groupings can be accounted for by the larger total number of conceptual items failing item analytical and EFA procedures and thus being eliminated from the final item pool. This changed the characteristics of the total item pool in terms of semantics and therefore the wording of the final three domains had to be carefully considered so that each domain described the items that comprised it as accurately as possible. The greater methodological considerations regarding the adequate representation of the data generated at focus group level by the final scale is discussed in more detail in section 7.5 of the current chapter. What follows here is a critical exploration of how the conceptual framework and domains of the final scale can provide us with insights into what the ER-QLS measures.

As was mentioned previously, rather than a generic or psychological evaluation outcome measure, at every phase of research the ER-QLS was designed with the assessment of exercise-related life-quality in mind. A bottom up approach to item generation meant that participants representative of the population for whom the measure is intended acted as the primary source for shaping conceptual understandings of what indicators represent exercise-related life-quality. Furthermore, the item construction phase of research employed a careful and critical approach to ensuring that only those items considered as causal or effect indicators of life-quality were included in the test-item pool (see Chapter III). These procedures were purposefully aimed at shaping the final scale into a measure of life-quality rather than something else. The convergent validity analysis of the ER-QLS that was conducted at both total score and sub-domain level provided valuable insights into what indicators of life-quality are measured by each of the domains (i.e., mental and physical well-being, injury pain and illness and physical activity facilitators).

In summary, these correlation analyses suggested that the mental and physical well-being domain of the ER-QLS measures those aspects of QoL that are similar to the

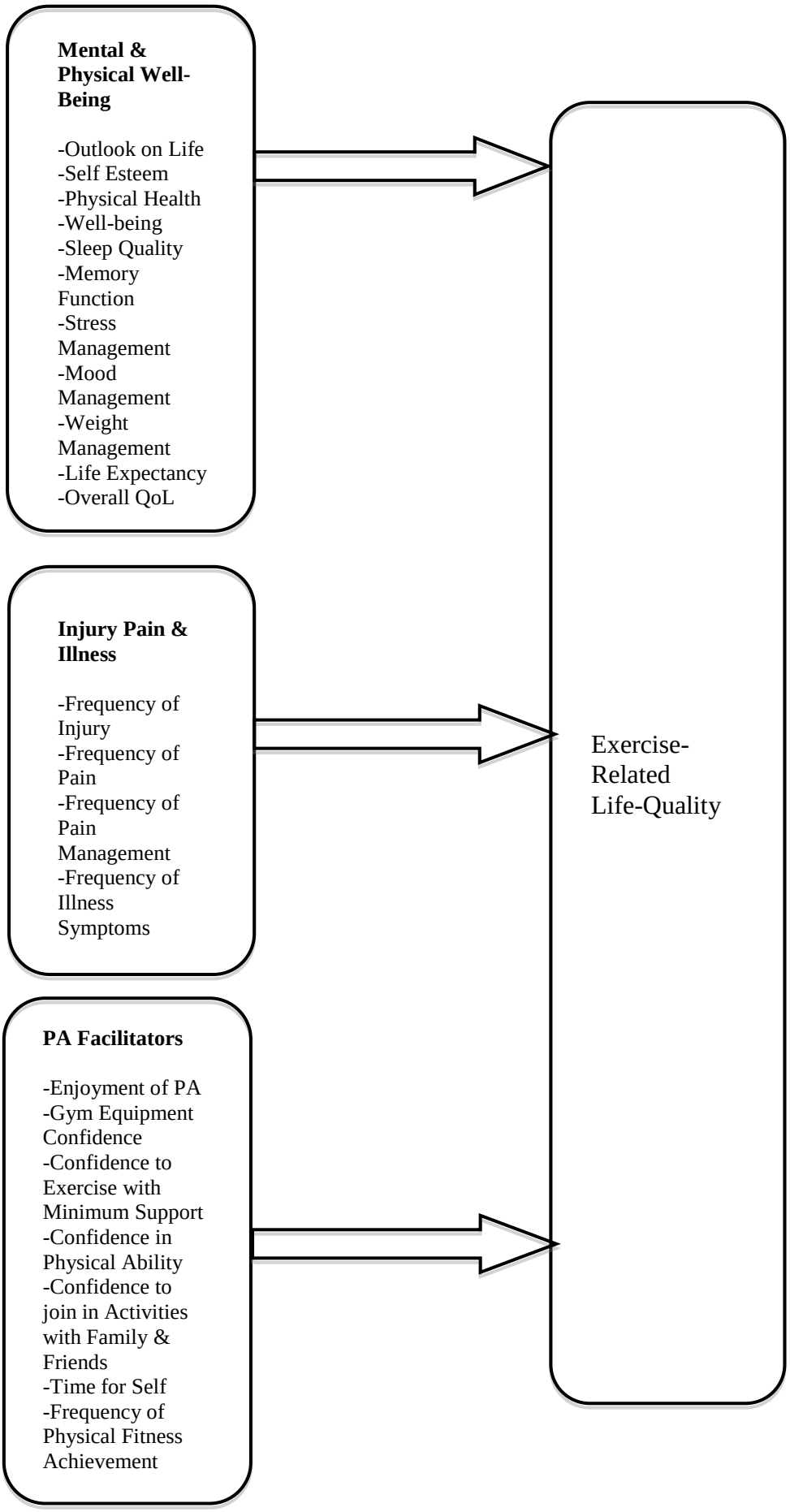
indices addressed by more generic measures (e.g., WHOQOL-BREF), but because of the procedures undertaken to formulate the items that comprise this scale, it is likely that this domain is representative of indices that are most relevant to exercise referral populations. Convergent validity analysis also suggested that the injury pain and illness domain and the PA facilitators domain target those indices of life-quality that remove the ER-QLS scale further from more generic measures. Specifically, convergent validity analysis suggested that the injury pain and illness domain of the ER-QLS targets physical indicators of life-quality and the PA facilitators domain targets primarily physical but also psychological aspects of life-quality. However the wording of items contained within these domains results in a measure that is more suited to exercise referral. For example, one of the items grouped into the injury pain and illness domain asks "how much does any injury you may have prevent you from being physically active?" One of the items grouped into the PA facilitators domain asks "how confident are you that you can exercise in a leisure centre with minimum support?"

Chapter I explained that within exercise referral settings, two of the most commonly used generic measures of life-quality are the SF-36 (Brazier et al., 1992) and the EQ-5D (Brooks, 1996). The eight dimensions of the SF-36 include: vitality, physical functioning, bodily pain, general health perceptions, physical role functioning, emotional role functioning, social role functioning and mental health. While the five dimensions of the EQ-5D are: mobility, self-care, usual activities, pain/discomfort and anxiety/depression. The domains that comprise the generic QoL measure, the WHOQOL-BREF (Skevington et al., 2004) that was used to demonstrate the convergent validity of the ER-QLS are: physical health, psychological health, social relationships and the environment.

The domains for each of these three measures are not too dissimilar to the ER-QLS (mental and physical well-being, injury pain and illness and physical activity facilitators). However, the key difference is that some of the items contained within the ER-QLS specifically refer to physical activity and exercise for example, "how much do you currently enjoy physical activity?" and "how confident are you using gym equipment?" These questions are asked alongside more generic questions relating to sleep quality, memory function and overall QoL, for example that were reported by focus group participants as having particular significance to their exercise related QoL. The ER-QLS also includes a domain that specifically addresses

those indicators that facilitate PA. In this respect, conceptual and measurement challenges aside, if the SF-36, EQ-5D and WHOQOL-BREF purport to measure generic QoL and the key difference between these measures and the ER-QLS is the specificity of exercise and physical activity (demonstrated by item wording and PA facilitators domain) then it is reasonable to propose at this stage of the scale's development that the ER-QLS is a measure of exercise-related life-quality. What the ER-QLS can measure that more generic measures cannot or may be limited by therefore are those specific aspects of life-quality that are responsible for perceptions of exercise-related life-quality. Much in the same way that health-related QoL measures are designed to target those indicators that are particularly responsible for the ill-health condition for which they have been constructed (e.g., Doward et al., 2003; Las Hayas et al., 2006). Figure 15 illustrates how the domains of the ER-QLS inform what is being measured when attempts are made to measure exercise-related life quality.

Figure 15. Measurable indicators of exercise-related life-quality



Further field testing is required to assess the performance of the 22-item ER-QLS (see 7.6.2) but in the interim, there is sufficient evidence to suggest that the measure is capable of detecting those indicators that a representative sample of exercise referral participants reported as important to their perceptions of exercise-related life-quality. As mentioned earlier, and presented in more detail in Chapter II, there were some reports made by focus group participants that did not transfer from the focus group phase of research into item construction and psychometric analyses because they were either not causal or effect indicators of life-quality, too vague in their description or used jargon language. The implications of the omission of these items is discussed in the following section.

7.4 Beyond Cause and Effect

Allison et al. (1997) have argued that QoL is a dynamic construct that may result in changes in the standards by which a respondent to a QoL measure may assess his/her QoL. This can prove particularly problematic when attempting to assess change over time. In this respect an assessment of those factors that contribute to a respondent's perception of life-quality at the time of assessment becomes critical. Furthermore, Hunt (1997) has advised that the measurement of QoL for individuals within medical settings runs the risk of isolating them from the social and material conditions within which they exist. Hunt (1997) goes on to suggest that it is unlikely that a person's QoL will be affected if they return from hospital to an unheated home with no employment, money and a dependant relative, for example. Again, such considerations make it difficult for scale developers and those with a responsible approach to health evaluation to suggest that a single approach to measurement is able to capture these lifestyle factors adequately.

The term 'self-efficacy' was considered problematic during item construction because of the use of jargon. Additionally, the terms 'feeling better' and 'self-awareness' were also considered problematic for item construction because the terms were considered too vague. All other areas of exercise-related life-quality identified as important by focus group participants were omitted from the initial item construction phases of research because they were considered neither causal nor effect indicators of life-quality (e.g., Fayers et al., 1997). These were: exercise referral to initiate PA, assistance with surgery and good instructors as identified by

their knowledge level, support, communication and level of familiarity (Chapter II; Table 8). These categories are perhaps better considered as process-related aspects of exercise-related life-quality. Omitting an exploration of the significance of these reported categories for exercise referral and the measurement of exercise-related life-quality would not only be unscientific but in terms of the participants included in the focus group phase of research, arguably also unethical (Hunt, 1997). Therefore, what follows is an examination of these process-related omitted categories with respect to how they may be included within assessments of exercise-related life-quality.

Focus group participants reported that exercise referral had played an important role in initiating their PA behaviour. The notion of exercise behaviour change has received a considerable amount of attention within the literature from both a theoretical (e.g., Ajzen & Fishbein, 1980; Bandura, 1988; Deci & Ryan, 1985, 2000; Heinzelmann & Bagley, 1970; Miller & Dollard, 1941; Prochaska, DiClemente, & Norcross, 1992) and also an applied perspective in terms of behaviour change counselling (BCC) within exercise settings (e.g., Breckon, Hally Johnston & Hutchinson, 2008; Hilton & Poulter, 2009; Laitakari & Asikainen, 1998; Nupponen, 1998.) It is this applied perspective that is of particular relevance to the current research and those process categories identified as important by focus group participants but omitted from the ER-QLS. Not only is this relevant from a stand-alone initiation of exercise behaviour perspective, but because a BCC approach suggests a certain level of person-centredness (e.g., Hilton & Poulter, 2009). As a consequence, it is likely that this approach would facilitate the identification of additional factors that are important to the client referred for exercise. Such an approach may well elicit information regarding those omitted focus group reports and/or others.

Equally, this approach may facilitate the reporting of items that were rejected during psychometric procedures. For example, lifestyle factors critical to the client's maintenance of exercise such as the barriers reported by focus group participants in the current study (i.e., time, transport, poor health, childcare and the weather) and similar barriers also reported previously (Biddle & Mutrie, 2001). Subsequently, this would allow the exercise instructor to collaboratively develop barrier-reducing strategies with clients to enable them to maintain their exercise behaviour.

It is likely that a BCC approach would facilitate a rapport between the exercise instructor and client (e.g., Miller & Rollnick, 2002) which has been deemed as important for questionnaire administration (e.g., Rust & Golombok, 2009) and that such a rapport may enhance client perceptions of instructor support, communication and familiarity which were reported as important omitted process categories by focus group participants. Arguably, such levels of rapport may well reduce the possibility of socially desirable responses to the ER-QLS or any other measure used in practice (e.g., Morten, 2010). In addition, although the frequency of missing data for the ER-QLS did not exceed the level of 4% suggested by Fayers and Machin (2000) it is reasonable to suggest that for the same reasons a BCC approach may facilitate low levels of missing data.

In terms of a BCC-consistent approach, the importance of instructor knowledge (an omitted process related focus group outcome) at first glance may be more tricky to incorporate. For example, motivational interviewing (MI) approaches that are typical of BCC strategies within clinical exercise settings (e.g., DH, 2009b ; Hilton et al., 2009) suggests that the client is the expert in their own behaviour change process. In this respect, the way in which information is provided by the exercise instructor is a delicate process that requires skill in order to avoid being too direct or to avoid what MI terms the 'righting-reflex' (Miller & Rollnick, 2002). However, these BCC approaches would not limit the instructor in providing exercise-related knowledge should the client require it as such approaches recognise the necessity to switch between differing communication styles for the benefit of the client (e.g., Rollnick, Miller & Butler, 2008). Furthermore, BCC has the capacity to be integrated into the administration of the ER-QLS in any of the three methods (i.e., self-completed, interview or telephone interview). Certainly cognitive pre-testing (Chapter III) revealed that respondents were keen to elaborate upon the responses they gave, providing more context and background to their decisions under all three conditions. In practice, such participant dialogue would provide valued opportunities to consult with the client in a BCC consistent manner.

BCC approaches to exercise behaviour change have been coined as brief interventions and have received support for their development and delivery (e.g., NICE, 2006a). Brief interventions have relevance for the current research for two primary reasons. Firstly, comparatively more support from NICE (2006a) for their

inclusion within clinical exercise settings than for exercise referral has given rise to an increase in their delivery (e.g., DH, 2009b) and it may be that there are lessons to be learned from how this approach may be combined with that of exercise referral to enhance the initial and sustained exercise behaviour change of those referred. Combining these approaches to behaviour change in such a way has previously been recognised as a method of improving exercise referral outcomes (Isaacs et al., 2007). The second reason summarises some of the key discussion points that have been presented in the current section of this thesis and that is that the learning from brief interventions may have direct implications for how the ER-QLS should be administered. As mentioned earlier, the importance of establishing and maintaining rapport for psychometric testing has previously been recognised (e.g., Rust & Golombok, 2009) and a BCC or brief intervention approach to the administration of the ER-QLS may well facilitate the accuracy of measurement through a reduction in socially desirable responses (e.g., Morten, 2010) and frequency of missing data, for example. Furthermore, such an approach may also limit the difficulties inherent in the measurement of a dynamic, shifting construct such as life-quality as identified by Allison, (1997) because working in such a way would allow for a greater exploration of those lifestyle factors that may affect the reference point a client uses to assess their life-quality. Additionally, combining these approaches may provide valuable opportunities to address the dangers of isolating respondents from the social and material conditions within which they exist (Hunt, 1997) which were described at the beginning of the current section of this chapter.

7.5 Strengths and Limitations of the Research Programme

The holistic evaluation of exercise referral schemes is currently inadequate and one of the challenges of integrating the psychological evaluation of exercise referral into routine practice, in particular life-quality is the lack of a population-specific instrument with which to capture this valuable data. Therefore, at the macro level one of the main strengths of the thesis is the nature of its applied focus - developed as a direct response to experiences in the field and designed to contribute to solving a common evaluation difficulty. The ER-QLS is the first exercise-referral sensitive measure of QoL to have been developed.

The mixed methods (i.e., both qualitative and quantitative) approach to achieving the research outcome and the aims of each study should be considered a particular strength of the research programme. The careful selection of methods at each stage of the research ensured that the most appropriate approach was applied to fulfilling the specific study aims. Furthermore, the use of reflective diaries and recording of process-related information either using NVivo software or by hand ensured that the appropriateness and methodological rigour of the series of studies that comprised the complete thesis were given considerable attention. This mixed-method approach to answering a research question and solving a problem for practice is particularly desirable as the approach is not bound to a single paradigm (e.g., Silverman, 2000).

Similarly, the inclusion of researcher reflexivity and audit trails during the qualitative phases of research (focus groups and cognitive pre-testing) have enhanced the integrity, rigour and trustworthiness of the qualitative research (Finlay, 2002; Koch, 1994; Webb, 1992, 1996). Iterative phases of peer de-briefing (Spall & Stephen, 1998) were a key consideration throughout the research programme. This was introduced at important stages of data interpretation during both the qualitative and quantitative phases of research and such procedures have been noted as contributing to the credibility of qualitative research in particular (Lincoln & Guba, 1985).

As the title suggests, qualitative approaches to research are not concerned with sample size as are quantitative approaches. Achieving adequate power for a statistical procedure or the importance of generalizability of results is of no relevance. However, given that the primary function of the focus groups (Chapter II) was to generate items for a test measure, conducting the groups until the point of data saturation (Calder 1977; Holloway, 1997) ensured that the QoL related issues of importance that were reported by participants were highly representative. Even at these early stages of the research process, this approach reduced the likelihood of a final scale that would be too specific (Klein, 2000, Cattell, 1973) and unable to capture the breadth of measurement required from a rigorous health measurement instrument. Similarly, the health conditions with which participants were referred for the focus groups (Chapter II) and psychometric analyses (Chapters IV and V) were representative of those with which individuals are usually referred into an exercise scheme and therefore has further enhanced the likelihood of a measure that is particularly relevant for the population for whom it is intended. Furthermore, at each

step of the research process the ER-QLS was designed with QoL definition and measurement issues in mind. This was particularly relevant by the inclusion of items that were representative of either causal or effect items of QoL (Fayers et al., 1997). These considerations have added to the robustness of the final measure.

A final recognition of the strengths of the research programme is that the resultant ER-QLS has been designed for flexible use, the measure may be administered by self-complete, interview or telephone methods and therefore enhances the likelihood that it will be used in practice. Similarly, the flexibility of the new scale's scoring method (i.e., either as a single score or sub-dimensionally) will enhance the variety of it's application within exercise referral settings as an evaluation tool and/or for the purposes of research.

One of the limitations of the current research is inherent in most initial scale development and validation studies in that the resultant scales are typically not validated to detect change over time (responsiveness). This is true of the current research. The known group construct validity has provided promising insights into the potential responsiveness of the measure although it is acknowledged that this data reflects different groups. There are plans for this research to be undertaken as additional longitudinal work which is typical for new scale development (e.g., Congleton, Hodson & Skingle, 1998).

The participants included in the focus groups were recruited from a single local scheme. It's possible that recruiting from others may have benefitted this stage of research in terms of enhancing the representativeness of the participant sample. However, previous focus group studies have focussed on a single scheme (e.g., Crone-Grant & Smith, 1999; 2001) and cognitive pre-testing that involved two, albeit local schemes and the larger scale cross-England testing of the measure (phases three and four of scale construction) did not highlight and difficulties in this respect.

The scoring protocol of the ER-QLS makes use of percentiles to assist with providing some context as to the meaning of respondent total scores. While not a limitation of the research as such (again initial scale development research tends to utilise the same system because only so much can be achieved during initial design and validation studies), future development of the ER-QLS may benefit from

revisiting these percentiles using a much larger cohort in order to establish more robust norms.

Finally, although the ER-QLS is designed for use in the UK, it is recognised that due to practical limitations outlined in Chapter IV only those schemes based in England were available to participate in the research at the time of scale testing. The ER-QLS was designed with the protocols for exercise referral delivery described in the NQAF (DH, 2002) in mind (Chapter I). These guidelines are designed to reflect the design and delivery of schemes operating across the UK. As such it is anticipated that the lack of inclusion of schemes based in Scotland, Wales and Northern Ireland for the current study will not impede on the use of the ER-QLS in these areas because the design and implementation of schemes based here reflect those included in the current study. There are plans to test the final 22-item version of the ER-QLS which will provide opportunities for additional schemes outside of England to be included in the scales future development. This will also provide a valuable opportunity to assess if the final scale is adequate for use across the UK.

7.6 Future Research Directions

The research undertaken for the current thesis has a number of potential future research directions. These are outlined below.

7.6.1 Qualitative Research and Theoretical Contributions

As mentioned previously, the primary function of the focus group data were for item generation and the level of analysis employed was suitable for this aim. However, future study could re-visit the qualitative data to explore the relationships between the coded data at a more thorough level. Such analysis may also be conducted from a theory driven perspective. Perhaps to further our understanding regarding theories of behaviour change. Some of the data regarding the participants relationship with the exercise instructor, for example may help to explore Edmunds et al's. (2006) SDT focussed findings in terms of how perceptions of autonomy support provided by exercise class leaders may predict psychological need satisfaction.

The qualitative phase of data analysis also provided some initial insights into what aspects of exercise-related life quality were particularly important to participants at the start of their referral programme, during the programme and upon completion. There was also some overlap between items of importance at each of these three stages and the role and attributes of the exercise instructor (i.e., knowledge, support, communication and familiarity) were reported as important to the referral experience across all three stages. Additional qualitative research could be undertaken to extend these initial observations and to assess this stage-matched proposal more thoroughly. Such work would require recruiting exercise referral participants who are at each stage of their programme (i.e., start, mid-referral and completion) into focus groups. This would ensure adequate representation of reports made at each stage. Additionally, interview guides that are specifically designed to capture participant reports regarding the role and importance of exercise-related life quality in relation to these stages of referral would need to be developed.

Similarly, comparatively little is known about the reasons why some of those referred for exercise fail to complete their initial programme of exercise (usually around 12 weeks) than the benefits of those who do attend. Attrition rates are a critical consideration for physical activity interventions. For example, reducing attrition in physical activity programmes for older adults has previously been investigated. Jancey, Lee, Howat, Clarke, Wang and Shilton (2007) recruited 248 insufficiently active older adults (65-74 years) into a local exercise programme comprising walking and strength exercises. Logistic regression analyses indicated that those lost to attrition (n = 86; 35%) resided in areas of lower socioeconomic status, were overweight, less physically active and had lower walking and self-efficacy scores and higher loneliness scores compared to completers (n = 162; 65%). Exit interviews with eight completers and eight who dropped out were conducted and although these qualitative reports are documented in comparatively less detail than the regression analysis, the qualitative data offers support for the quantitative data. Jancey et al. (2007) suggested that the early assessment of these characteristics could help to improve the retention of individuals participating in such programmes.

Adopting a purely qualitative or mixed methods approach similar to the study conducted by Jancey et al. (2007) to explore the reasons for attrition in exercise referral schemes would provide invaluable insights to the development of schemes in the future. Such a study would contribute to evaluating why some schemes may be

more successful than others, honour proposed guidelines regarding the appropriate evaluation of exercise referral schemes in the UK (NICE, 2006a) and contribute to the development of related public health strategies and interventions that aim to reduce sedentary behaviour.

Future research may also wish to investigate the conceptual framework of exercise-related life-quality (Figure 8) that was developed from the focus group phase of the current research to determine if it can be used to explain the exercise referral-QoL relationship in more detail. There is a wealth of research that has demonstrated improvements in QoL for various clinical populations following a programme of exercise (e.g., Burnham & Wilcox, 2002; Milne et al., 2007; Petajan et al. 1996) and researchers have scope to employ latent concepts such as self-efficacy for example, to explain the exercise-QoL relationship (e.g., Motl & Snook, 2008). Additionally, there are a number of health behaviour and behaviour change theories available to support research that aims to assess how individuals may engage in exercise long enough to experience improvements in life-quality (e.g., Ajzen & Fishbein, 1980; Bandura, 1988; Deci & Ryan, 1985, 2000; Heinzelmann & Bagley, 1970; Miller & Dollard, 1941; Prochaska, DiClemente, & Norcross, 1992). However, to date a conceptual framework of exercise-related life-quality is unavailable. There is a valuable opportunity to match the outcomes of exercise-related life-quality studies (both qualitative and quantitative) against the contents of the framework (i.e., leisure facilities and lifestyle physical activity, discomfort and fitness, mental and physical well-being and confidence and motivation) to further assess the appropriateness of the framework and contribute to theoretical understandings of exercise-related life-quality.

7.6.2 Further Validation of the ER-QLS and Longitudinal Study

As mentioned previously, the ER-QLS has yet to be administered in its 22-item form and further testing will be required to assess the performance of the final scale. Future development of the ER-QLS may include a longitudinal examination of the scales performance to ascertain how well the measure can detect change over time (responsiveness). Utilising a larger cohort to establish norms for percentile scores

and threshold descriptions would also comprise part of the longitudinal work and would further inform the scoring protocol.

Undertaking additional confirmatory factor analytical research for newly developed scales is commonplace. Additional CFA using larger sample sizes would be particularly beneficial to the ER-QLS not only as part of analysis that is typical of that which follows the procedures undertaken for initial scale development but especially because CFA would allow for an examination of whether the three-factor or single factor model of the ER-QLS is the most appropriate. This analysis would require engagement from exercise referral schemes from across the UK in much the same way as that reported in the current research which has two added valuable research opportunities. Firstly, as mentioned earlier this will provide those schemes not involved in the current phases of scale construction research to be included in these additional studies and secondly, the face validity of the 22-item version of the ER-QLS could be assessed from both a patient and a practitioner perspective.

If the popularity of the measure grows in practice, it may also be necessary to translate the scale into other languages to reflect the cultural diversity of the UK. This is a procedure that is typical of future research regarding the progression of new scale development (e.g., Montazeri, Goshtasebi, Vahdaninia, & Gandek, 2005). The procedures required to undertake such research would mirror those employed by the current thesis. Again, the process of cognitive interviewing would be particularly important in order to ascertain if there were any translation problems with items and the utility of the scale cross-culturally.

The ER-QLS is easy to administer using either of the three methods that were cognitively pre-tested (i.e., self-complete, interview or telephone) and as such no special training is required for exercise referral professionals to be able to use the measure in practice. However, one of the recommendations made in this thesis is that BCC/BI techniques are employed to administration in order to ensure that the process, causal and effect indicators of exercise-related life-quality are adequately addressed during consultations with clients. In this respect, exercise referral professionals who haven't been in receipt of such training would benefit from attendance at training such as that delivered as part of the DH's Let's Get Moving (LGM) programme (DH, 2009b). This training has been developed to equip exercise professionals with the necessary skills to deliver a BI including MI. Indeed, the ER-

QLS has the capacity to complement the LGM programme more generally as one of the opportunities for exercise offered as part of the programme design is exercise referral.

7.7 Implications for Policy and Practice

The integration of the ER-QLS into routine evaluation has direct implications for public health practice and future policy as the measure may help to address some of the psychological outcomes of exercise referral that have been identified as the most likely health parameters to change (DH, 2002) but that are commonly overlooked (Dugdill et al., 2005). There is potential for the narrow focus of efficacy reviews such as the NICE (2006a) review regarding only an assessment of PA levels to include assessments of exercise-related life-quality into evidence-based reviews of exercise referral. The health benefits of PA are so diverse (e.g., mental, physical, social, economic and environmental; e.g., DH, 2004b) that it only makes sense to evaluate holistically. Not doing so runs the risk of missing valuable outcome data which has been recognised previously (Dugdill et al., 2005 Dugdill & Graham, 2005). Furthermore, overlooking the potential responses to exercise for those referred could also be deemed as unethical (e.g., Hunt, 1997).

The qualitative phase of research undertaken for the current thesis has highlighted the value of using focus groups to critically explore issues of evaluative importance to exercise referral schemes. Combining the administration of the ER-QLS with such qualitative approaches as part of routine evaluation practice within exercise referral settings may well assist with gaining a more balanced and informed perspective regarding their ability to impact upon the health parameters and PA behaviour of those referred. Indeed Woods, Agarwal, Young, Jones and Sutton (2004) have published a comprehensive guide to the use of integrative approaches to qualitative and quantitative evidence within healthcare settings. What the authors fail to recognise however, and what is often a common misconception in practice is that the successful implementation of qualitative evaluation methods at practitioner level requires specific training and/or opportunities for skill acquisition. Conducting focus groups or individual interviews, for example in a manner that is most likely to yield rich meaningful data is a skill that must be practiced and acquired over time. The

recommendation for the ER-QLS to be administered in a BCC/BI consistent manner may help to bridge this gap if exercise referral professionals wish to evaluate their scheme qualitatively, have not received training and cannot afford to pay for an academic institution to undertake the work on their behalf. Specifically, because BCC skills have much in common with what is deemed good qualitative interviewing practice (e.g., Patton, 1990).

Although BCC/BI focussed initiatives to change PA behaviours at the individual level have received support (e.g., NICE, 2006a) and a favourable evaluation of a UK BI feasibility trial has resulted in a national roll-out of the design (DH, 2009b) there are no recommendations regarding the most effective or feasible approach to BCC for professionals working in environments that require supporting individuals to change PA behaviour. The growing evidence base for MI to successfully support a number of health behaviour changes that are relevant to those who are inactive or commonly referred for exercise for example, cardiovascular health and hypertension (Brodie & Inoue, 2005) and diet and lipids (Richards, Kettelman & Ren, 2006) has lead to an increased interest in the method.

As mentioned earlier, training based upon the MI method has been developed to facilitate exercise professionals ability to work in a BI consistent manner (e.g., Let's Get Moving; DH, 2009). The implications of recommending that the ER-QLS is administered in the same way has direct implications for practice and the future development of exercise referral in the following ways. Sharing approaches to patient care in this respect enhances the continuity of patient experiences throughout their referral pathway. Something that has been referred to as continuity of care in general practice (Freeman & Hjortdahl, 1997). BI training has already been developed therefore, this reduces any potential barriers that exercise referral professionals would have to it for the purposes of ER-QLS administration. The integration of the ER-QLS into practice will further support this BI work which is very much in its infancy and an evaluation of practitioner's experiences of incorporating the MI approach into consultations may help to add some clarity as to the suitability of the method within exercise settings. In the longer term this may inform the development of best practice guidelines which has the potential to impact upon health policy.

Physical inactivity is a complex problem to which there is no simple solution and those responsible for transport, leisure services and the environment, for example all have a role to play in reducing health inequalities and enhancing QoL (Hilton, 2010). In a world where policy is shaped by evidence-based practice, ineffective evaluation strategies could impact negatively upon public health policy. Taking into consideration the expertise of the health professional in addition to the values and preferences of the client has been identified as central to evidence-based practice (Mazurek Melnyk & Fineout-Overholt, 2005). Using strategies that research suggests are the best approach to achieving agreed aims, changing practice and a systematic appraisal of the best available evidence have also been attributed to an evidence-based practice approach (Fleming, 2009). However, with the ever increasing importance of demonstrating that what is being done within healthcare settings is effective, it may be more appropriate to adjust the notion of evidence-based practice to focus instead upon practice-based evidence. Contributions to the range of process and outcome parameters that are incorporated into such evaluations, perhaps particularly in terms of the tools available to practitioners and access to appropriate training may facilitate this aim. In this respect, it is expected that the integration of the ER-QLS into the routine evaluation of exercise referral schemes will contribute to this recent development in public health.

7.8 Conclusions

Health and exercise practitioners have a growing interest in capturing the more holistic outcomes of varied interventions designed to tackle health inequalities and reduce disease. Consequently, measures of life-quality have been deemed as "the essence of healthcare" (Holmes, 2005, p. 493). Alongside physiological indices such measures have a key role to play in monitoring health outcomes at client/patient level. However, generic QoL may be more difficult to measure than disease or population specific QoL, especially given that there is little agreement regarding how to define the term (e.g., Holmes, 2005; Hunt, 1997). While disease or population specific measures attempt to navigate this difficulty, inadequate explanations of what, exactly these instruments do and do not measure and the lack of any theoretical or conceptual considerations make claims regarding the specificity and sensitivity of these QoL measures difficult to understand. The current thesis has attempted to

address these measurement challenges by providing a clear conceptual framework of exercise-related life-quality and a transparent description of what the ER-QLS purports to measure.

The ER-QLS is a reliable and valid measure of exercise-related life-quality that has been developed from a mixed methods (i.e., both qualitative and quantitative) research approach including purposeful methodologically tailored phases of research that have contributed to meeting the overarching research aim. While other approaches to the facilitation and evaluation of psychological processes and outcomes of exercise referral are acknowledged (e.g., BCC and the use of qualitative methods). The ER-QLS is intended to contribute to the tools currently available to practitioners.

References

Aaronson N. K. (1990) Quality of life research in cancer clinical trials: A need for common rules and language. *Oncology* 4, 59-66.

Adair, C. E., Marcoux, G.C., Bischoff, T.F., Cram, B.S., Ewashen, C.J., & Pinzon, J…Brown, K. E. (2010). Responsiveness of the Eating Disorders Quality of Life Scale (EDQLS) in a longitudinal multi-site sample. *Health and Quality of Life Outcomes,* 8(83), 2-11.

Aleamoni, L. M. (1976). The relation of sample size to the number of variables in using factor analysis techniques. *Educational and Psychological Measurement,* 36, 879-883.

Allison, P. J**,** Locker, D., & Feine, J. S. (1997). Quality of life: A dynamic construct. *Social Science and Medicine,* 45(2), 221-230.

Alexander, J. L., & Wagner, C. L. (2006). How cardiac rehabilitation relates to quality of life. *Rehabilitation Nursing*, 31(4), 155-165.

Amrhein, P. C., Miller, W. R., Yahne, C. E., Palmer, M., & Fulcher, L. (2003). Client commitment language during motivational interviewing predicts drug use outcomes. *Journal of Consulting and Clinical Psychology,* 71, 862-878.

Andrews, F. M. & Withey, S. B. (1976). *Social indicators of well-being: American's perceptions of life-quality*. New York: Plenum Press.

Andrich, D. (1988). *Rasch models for measurement*. Beverly Hills: Sage.

Andrich, D. (2004). Controversy and the Rasch model: A Characteristic of Incompatible Paradigms? *Medical Care,* 42, 1-16.

Angermeyer, M. C., Holzinger, A., Kilian, R., & Matschinger, H. (2001). Quality of life- as defined by schizophrenic patients and psychiatrists. *International Journal of Social Psychiatry,* 47(2), 34-42.

Ajzen, I., & Fishbein, M. (1980). Understanding attitudes and predicting social behaviour. *Englewood Cliffs, N.J: Prentice-Hall.*

Badia, X., Arribas, F., Ormaetxe, J. M., Peinado, R., & Terreros, M. S. (2009). Development of a questionnaire to measure health-related quality of life (HRQoL) in patients with atrial fibrillation. *Health and Quality of Life Outcomes,* 5(37), 1-8.

Bandura, A. (1988). Organizational application of social cognitive theory. *Australian Journal of Management,* 13(2), 275-302.

Barrett, P. T., & Kline, P. (1981). The observation to variable ratio in factor analysis. *Personality Study and Group Behavior,* 1, 23-33.

Basch C. E. (1987). Focus group interview: An underutilized research technique for improving theory and practice in health education. *Health Education Quarterly*, 14,(4), 411–448.

Bates, B. R., Lynch, J. A., Bevan, J. L., & Condit, C. M. (2004). Warranted concerns, warranted outlooks: A focus group study of public understandings of genetic research. *Social Science and Medicine,* 60(2), 331-344.

Bauman, A.E., Fardy, J. H., & Harris, P. G. (2003). Why bother with patient-centred care? *Medical Journalists Association*, 179, 253-256.

Beaty, P. C., & Willis, G. B. (2002). Research synthesis: The practice of cognitive interviewing. *Public Opinion Quarterly*, 71(2), 287-311.

Belson, W. A. (1981). *The design and understanding of survey questions.* Aldershot: Gower.

Bennett-Levy, J. (2003). Reflection: A blind spot in psychology, *Clinical Psychology,* 27, 16-19.

Berg, B. L. (2004). *Qualitative research methods for the social sciences.* Boston: Pearson.

Bergner, M., Bobbitt, R. A., Kressel, S., Pollard, W. E., Gilson B. S., & Morris, J. R. (1976). The sickness impact profile: Conceptual formulation and methodology for the development of a health status measure. *International Journal of Health Services*, 6(3), 393–415.

Biddle, S. J. H. (2000). Emotion, mood and physical activity. In S. J. H. Biddle, K. R. Fox, and S. H. Boutcher (Eds.), *Physical Activity and Psychological Well-being* (pp. 63–87). London: Routledge.

Biddle, S. J. H., & Mutrie, N. (2001). *Psychology of physical activity: Determinants, well-being and interventions.* London: Routledge.

Blaikie, N. W. H. (1993). *Approaches to social enquiry.* Cambridge: Polity Press.

Blair, J., & Presser, S. (1993). Survey procedures for conducting cognitive interviews to pretest questionnaires: A review of theory and practice. *Proceedings of the Section on Survey Research Methods of the American Statistical Association,* 370-375.

Bloodworth, A., &, McNamee, M. (2007). Conceptions of well-being in psychology and exercise psychology research: A philosophical critique. *Health Care Analysis*, 15, 107-21.

Booth, M. L. (2000). Assessment of physical activity: An international perspective. *Research Quarterly for Exercise and Sport,* 71(2), 114-20.

Booth, M. L., Bauman, A. Owen, N., & Gore, C. J. (1997). Physical activity preferences, preferred sources of assistance, and perceived barriers to increased physical activity among physically inactive Australians. *Preventative Medicine*, 26(1), 131-137.

Bowden, D., Fesinmeyer, M. D., Yasui, Y., Tworoger, S., Ulrich, C. M., & Irwin, M…McTiernan, A. (2006). Randomised trial of sedentary middle aged women: Effects on quality of life. *International Journal of Behavioural Nutrition and Physical Activity,* 3, 1-9.

Bowling, A. (1997*) Research methods in health.* Buckingham: Open University Press.

Brace, I. (2004). Objectives in writing a questionnaire. In I. Brace (Ed.), *Questionnaire design how to plan, structure and write survey material for effective market research* (pp. 7-22). London: Kogan Page.

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

Breckon, J. D., Halley Johnston, L., & Hutchison, A. (2008). Physical activity counselling content and competency: A systematic review. *Journal of Physical Activity and Health*, 5, 398-417.

Brod, M., Stewart, A. L., Sands, L., & Walton, P (1999). Conceptualization and measurement of quality of life in dementia: The dementia quality of life instrument (DQoL). *Gerontologist*, 39(1), 25-35.

Brodie, D.A., and Inoue, A. (2005). Motivational Interviewing to Promote Physical Activity for people with Chronic Heart Failure. *Journal of Advanced Nursing*, 50, 518-527.

Brooks, R., & EuroQoL Group (1996). EuroQoL: The current state of play. *Health Policy*, 37(1), 53-72.

Brown, J., Bowling, A., & Flynn, T. (2004). Models of quality of life: a taxonomy, overview and systematic review of the literature. London: European Forum on Population Ageing Research.

Bruce, B., Fries, J. F., Ambrosini, D., Lingala, B., Gandek, B., & Rose, M., & Ware, J. E. (2009). Better assessment of physical function: Item improvement is neglected but essential. *Arthritis Research and Therapy,* 11, 2-27.

Brunker, P. D., &, Brown, W. J. (2005). Is exercise good for you? *The Medical Journal of Australia,* 183, 538-541.

Bryant, F. B., & Yarnold, P. R. (1995). Principal-components analysis and confirmatory factor analysis. In L. G. Grimm & P. R. Yarnold (Eds.), *Reading and understanding multivariate statistics*. (pp. 99-137). Washington, DC: American Psychological Association.

Buck, D. S., Monteiro, F. M., Kneuper, S., Rochon, D., Clark, D. L., & Melillo, & Volk, R. J. (2005). Design and validation of the health professionals' attitudes toward the homeless inventory (HPATHI). *British Medical Council Medical Education*, 5(2), 1-8.

Bull, F. C., & Milton, K. E. (2010). A process evaluation of a "physical activity pathway" in the primary care setting. *BMC Public Health,* 10, 463-472.

Burckhardt, C. S., & Anderson, K. L. (2003). The quality of life scale (QOLS): Reliability validity and utilization. *Health and Quality of Life Outcomes,* 60(1), 22-29.

Burnham T. R., & Wilcox A. (2002). Effects of exercise on physiological and psychological variables in cancer survivors. *Medicine and Science in Sports and Exercise.* 34, 1863–1867.

Calder, B. J. (1977). Focus groups and the nature of qualitative marketing research. *Journal of Marketing Research,* 4, 353-364.

Calman K. C. (1984). Quality of life in cancer patients - a hypothesis. *Journal of Medical Ethics,* 10, 124-127.

Campbell, A. (1976). Subjective measures of well-being. *American Psychologist,* 31, 117-124.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin,* 56, 81-105.

Cardinal, B. J., & Sachs, M. L. (1996). Prospective analysis of stage-of exercise movement following mail-delivered, self-instructional exercise packets. *American Journal of Health Promotion,* 9, 430–432.

Carolan, M. (2003). Reflexivity: A personal journey through data collection. *Nurse Researcher,* 10(3), 7-14.

Carr, A. J., Gibson, B., & Robinson, P. G. (2001). Measuring quality of life: Is quality of life determined by expectations or experience? *British Medical Journal* 322(7296), 1240 –1243.

Carraher, S. M., & Buckley, M. R. (1995). The effect of retention rule on the number of components retained: The case of cognitive complexity and the PSQ. *Journal of Applied Psychology,* 81, 102-109.

Carron, A. V., Widmeyer, W. N., & Brawley, L. R. (1988). Group cohesion and individual adherence to physical activity. *Journal of Sport and Exercise Psychology*, 10(2), 127-138.

Caspersen, C. J., Powell, K. E., & Christenson, G. M. (1985). Physical activity, exercise and physical fitness: Definitions and distinctions for health-related research. *Public Health Reports*, 100, 126-131.

Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behaviour Research,* 1, 245-276.

Cattell, R. B. (1978). *The scientific use of factor analysis.* New York: Plenum.

Cella, D. F. (1994) Quality of life: Concepts and definition. *Journal of Pain and Symptom Management*, 9, 186-192.

Chalder, T., Berelowitz, G., Pawlikowska., Watts, L., Wessely, S., Wright, D. & Wallace, E. P. (1993). Development of a fatigue Scale. *Journal of Psychosomatic Research,* 37(2), 147-153.

Chandra, R. K. (2001). Whole health: A prescription for the new millennium. *Nutrition Research,* 21, 1-8.

Chin A Paw, M. J., Van Poppel, M. N., Twisk, J. W., & Van Mechelen, W. (2004). Effects of resistance and all round, functional training on quality of life, vitality, and depression of older adults living in long-term care facilities: a 'randomized' controlled trial. *BMC Geriatrics,* 4, 23-32.

Chong, L.Y., Chay, O. M., & Chuen, L. S. (2006). Is the childhood asthma questionnaire a good measure of health-related quality of life of asthmatic children in Asia? Validation among Paediatric Patients with Asthma in Singapore. *Pharmacoeconomics,* 24(6), 609-621.

Cohen S. R., & Mount, B. (1992). Quality of life assessment in terminal illness: Defining and measuring subjective well-being in the dying. *Journal of Palliative Care* 1992(8), 40-45.

Cohen, S. R., Mount, B. M., & MacDonald, N. (1996). Defining quality of life. *European Journal of Cancer,* 32(5), 753-754.

Collins, D. (2003). Pretesting survey instruments: An overview of cognitive methods. *Quality of Life Research,* 12, 229-238.

Comfrey, A. L., & Lee, H. B. (1992). *A First course in factor analysis*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Congleton, J. Hodson, M. E., & Skingle, F. D. (1998). Do Nottingham Health Profile scores change over time in cystic fibrosis? *Respiratory Medicine,* 92(2), 268-272.

Conrad, K. J., Wright, B. D., McKnight, P., McFall, M., Fontana, A., & Rosenheck, R. (2004). Comparing traditional and rasch analyses of the mississippi PTSD scale: revealing limitations of reverse-scored items. *Journal of Applied Measurement*, 5(1), 2-16.

Corbin, J., & Morse, J.M. (2003). The unstructured interactive interview: Issues of reciprocity and risks when dealing with sensitive topics. *Qualitative Enquiry*, 9(3), 335-354.

Costello, A. B., & Osborne, J. W. (2005). Best practices in exploratory factor analysis: four recommendations for getting the most from your analysis. *Practical Assessment, research and Evaluation,* 10(7), 225-234.

Courneya, K. S. (2005). Exercise and quality of life in cancer survivors. In G. Faulkner, & A. H. Taylor (Eds.), *Exercise, Health and Mental Health: Emerging Relationships* (pp. 114-134). London: Routledge.

Courneya K. S., Mackey J. R., Bell G. J., Jones L. W., Field C. J., & Fairey A. S. (2003).Randomized controlled trial of exercise training in postmenopausal breast cancer survivors: Cardiopulmonary and quality of life outcomes. *Journal of Clinical Oncology*, 21, 1660–1668.

Creswell, J. W. (1998). *Qualitative Inquiry and Research Design.* London: Sage.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52, 281-302.

Crone, D., Johnston, L., & Grant, T. (2004). Maintaining quality in exercise referral schemes: A case study of professional practice. *Primary Health Care Research and Development*, 5, 96-103.

Crone, D. Smith, A., & Gough, B. (2005). 'I feel totally at one, totally alive and totally happy': A psycho-social explanation of the physical activity and mental health relationship. *Health Education Research,* 20, 600-611.

Crone-Grant, D., & Smith, R. A. (1999). Broadening horizons. A qualitative inquiry on the experience of patients on an exercise prescription scheme. *Journal of Sports Sciences,* 17, 12.

Crone-Grant, D., & Smith, R. A. (2001). Exercise referral schemes in public and private leisure: A qualitative investigation. *Journal of Sports Sciences,* 19, 19.

Deci, E. L., & Ryan, R. M. (1980). The empirical exploration of intrinsic motivational processes. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (pp. 39–80). New York: Academic.

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. New York: Plenum.

Deci, E. L., & Ryan, R. M. (2000). The "what" and "why" of goal pursuits: Human needs and self determination of behaviour. *Psychology Inquiry,* 11, 227-269.

De Leeuw, E. D. (2001). Reducing missing data in surveys: An overview of methods. *Quality and Quantity,* 35, 147-160.

Denning, J. D., & Verschelden, C. (1993). Using focus groups in assessing training needs: Empowering child care workers. *Child Welfare League of America,* 72(6), 569-579.

Denzin, N. (1989). *The research act: A theoretical introduction to sociological methods* (3rd ed). New Jersey: Prentice Hall.

Denzin, N., & Lincoln, Y. (1994) *Handbook of qualitative research.* London: Sage.

Department of Health (2000). *Health survey for England, 1998.* London: The Stationery Office.

Department of Health (2002). Exercise referral systems: A national quality assurance framework. London: The Stationery Office.

Department of Health (2004a). *Choosing health. Making healthy choices easier*. London: The Stationery Office.

Department of Health (2004b). *At least 5 a Week. Evidence of the impact of physical activity and its relationship to health. A report from the Chief Medical Officer*. London: The Stationery Office.

Department of Health (2004c). *The NHS improvement plan: Putting people at the heart of public services.* London: The Stationery Office.

Department of Health (2005). *Choosing Activity. A Physical Activity Action Plan*. London: The Stationery Office.

Department of Health (2008). *Health survey for England, 2008.* London: The Stationery Office.

Department of Health (2009a). *Be active, be healthy: A plan for getting the nation moving.* London: The Stationery Office.

Department of Health (2009b). *Let's get moving feasibility study.* London: The Stationery Office.

De Silva, M. J., Harpham, T., Tuan, T., Bartolini, R., Penny, M. E., & Huttly, S. R. (2006). Psychometric and cognitive validation of a social capital measurement tool in Peru and Vietnam. *Social Science and Medicine,* 62, 941-953.

Diener, E. (2000). Subjective well-being the science of happiness and a proposal for a national index. *American Psychologist*, 55(1), 34-43.

Doward, L. C., Spoonberg, A., Cook, S. A., Whalley, D., Helliwell, P. S., & Kay, L…Chamberlain, M. (2003). Development of the ASQoL: A quality of life instrument specific to ankylosing spondylitis. *Annuals of the Rheumatic Diseases,* 62, 20-26.

Drennan, J. (2003). Cognitive interviewing: Verbal data in the design and pretesting of questionnaires. *Journal of Advanced Nursing,* 42, 57-63.

Dubbert, P.M. (2002). Physical activity and exercise: Recent advances and current challenges. *Journal of Consulting and Clinical Psychology,* 70(3), 526-536.

Dugdill, L., & Graham, R. C. (2005). Promoting physical activity: Building Sustainable interventions. In J. Gormley, & J. Hussey (Eds.) *Exercise Therapy: Prevention and Treatment of Disease* (pp. 240-255). Oxford: Wiley-Blackwell.

Dugdill, L., Graham, R., & Crone, D. (2009). Developing the evidence-base for physical activity interventions. In L. Dugdill, D. Crone &, R. Graham (Eds.), *Physical activity and health promotion: Evidence-based approaches to practice* (pp. 60 – 81). London: Wiley-Blackwell.

Dugdill. L., Graham, R. C., & McNair, F. (2005). Exercise referral: The public health panacea for physical activity promotion? A critical perspective of exercise referral schemes; their development and evaluation. *Ergonomics,* 48, 1390-1410.

Eddy, L., Khastou, L., Cook, K. F., & Amtmann (in press). Item selection in self-report measures for children and adolescents with disabilities: Lessons from cognitive interviews. *Journal of Pediatric Nursing.*

Edwards, S. D., Ngcobo, H. S. B., Edwards, D. J., & Palavar, K. (2005). Exploring the relationship between physical activity, psychological well-being and physical self perception in different exercise groups. *South African Journal for Research in Sport, Physical Education and Recreation,* 27(1), 75-90.

Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists.* Mahwah, New Jersey: Lawrence Erlbaum Associates Inc.

Engel, G. L. (1977). The need for a new medical model: A challenge for biomedicine. *Science,* 196, 129-136.

Engel, G. L. (1990). The essence of the biopsychosocial model: From 17th to 20th Century science. In H. Balner (Ed.), *a challenge for biomedicine?* (pp. 13-18). Amsterdam and Rockland: Swets and Zeitlinger, Inc.

Eysenck, H. J. (1953). *The structure of human personality.* London: Routledge.

Fayers, P. M., Hand, D. J., Bjordal, K., & Groenvold, M. (1997). Causal indicators in quality of life research. *Quality of Life Research,* 6, 393-406.

Fayers, P. M. & Machin, D. (2000). *Quality of Life: Assessment, Analysis and Interpretation.* Chichester: John Wiley & Sons.

Ferrans, C. E. (1996). Development of a conceptual model of quality of life. *Research and Theory for Nursing Practice*, 10(3), 293-304.

Fernandez, F. J. C., Garcia, M. T. M., Alvarez, C. R., Giron, M. J. I., & Aguirre-Jamie, A. (2007). Is there an association between physical exercise and the quality of life of hypertensive patients? *Scandinavian Journal of Medicine and Science in Sports*, 17, 348-355.

Fernandez-Lopez, J. A., Seigrist, J., Hernandez-Mejia, R., Broer, M., & Cueto-Espinar, A. (1994). A study of quality of life on rural hypertensive patients, comparison with the general population of the same environment. *Journal of Clinical Epidemiology*, 47, 1373- 1380.

Field, A. (2005). *Discovering statistics using SPSS* (2nd ed.). London: Sage.

Field, A. (2009). *Discovering statistics using SPSS* (3rd ed.). London: Sage.

Fischer, G. H., & Molenaar, I. W. (1995). *Rasch models: Foundations, recent developments and applications*. New York: Springer-Verlag.

Fischer, J. S., Rudick, R. A,. Cutter, G. R., & Reingold, S. C. (1999). The multiple sclerosis functional composite measure (MSFC): An integrated approach to MS clinical outcome assessment. *Multiple Sclerosis,* 5, 244-250.

Flanagan, J. C. (1978). A research approach to improving our quality of life. *American Psychologist*, 33, 138-147.

Flanagan, J. C. (1982). Measurement of the quality of life: Current state of the art. *Archives of Physical Medicine and Rehabilitation*, 3, 56-59.

Flemming, M. L., & Fitzgerald, G. (2009). Evidence based practice. In M.L. Fleming & E. Parker (Eds.), *Introduction to public health* (pp. 165-192). Australia: Elsevier.

Foddy, W. (1993). *Constructing questions for interviews and questionnaires. Theory and practice in social research.* London: Cambridge University Press.

Fogarty, M., & Happell, B. (2005). Exploring the benefits of an exercise program for people with schizophrenia: A qualitative study. *Issues in Mental Health Nursing*, 26, 341-351.

Fontana, A., & Frey, J. H. (2003). The interview: From structured questions to negotiated text. In N. Denzin, & T. Lincoln (Eds), *Collecting and interpreting qualitative aterials* (pp. 61-106) London: Sage.

Ford, J. K., MacCallum, R. C., & Tait, M. (1986). The application of exploratory factor analysis in applied psychology: A critical review and analysis. *Personnel Psychology*, 39, 291-314.

Fox, K. R. (1999). The influence of physical activity on mental well-being. *Public Health Nutrition,* 2, 411-418.

Fox, K. R. (2000). The effects of exercise on self-perceptions and self-esteem. In S. J. H. Biddle, K. R. Fox, & S. H. Boutcher (Eds.), *Physical activity and psychological well-being* (pp. 88–117). London: Routledge.

Fox, K. R., Biddle, S. J. H., Edmunds, L., Bowler, I., & Killoran, A. (1997). Physical activity promotion through primary health care in England. *British Journal of General Practice,* 47, 367-369.

Freeman, G., & Hjortdahl, P. (1997). What future for continuity of care in general practice. *British Medical Journal,* 314, 1870-1874.

French, R. C., McLean, M., Williams, M., & Sutton, S. (2007). What do people think about when they answer theory of planned behaviour questionnaires? A 'think-aloud' study. *Journal of Health Psychology,* 12(4), 672-687.

Garrett, N. A., Brasure, M., Schmitz, K. H., Schultz, M. M., & Huber, M.R. (2004). Physical inactivity direct cost to a health plan. *American Journal of Preventative Medicine,* 27, 304-309.

Gerbing, D. W., Hamilton, J. G. (1996). Viability of exploratory factor analysis as a precursor to confirmatory factor analysis. *Structural Equation Modeling,* 3, 62–72.

Gianaros, P. J. Muth, E. R., Mordkof, J. T., Levine, M. E., & Stern, R. M. (2001). A questionnaire for the assessment of the multiple dimensions of motion sickness. *Aviation, Space and Environmental Medicine,* 72(2), 115-119.

Gidlow, C. & Murphy, R. (2009). Physical activity promotion in primary health care. In Dugdill, L., Crone, D., & Murphy, R. (eds). *Physical activity and health promotion. Evidence-based approaches to practice.* London: Wiley Blackwell.

Gilgun, J.F. (2002). Conjectures and refutations: Governmental funding and qualitative research. *Qualitative Social Work,* 1(3), 359-375.

Gorely, T., & Gordon, S. (1995). An examination of the transtheoretical model and exercise behaviour in older adults. *Journal of Sport and Exercise Psychology*, 17, 312–324.

Gorsuch, R. L. (1983). *Factor Analysis* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

Graham, R.C. Dugdill L., & Cable, N. T. (2005). Health Professionals' perspectives in exercise referral: Implications for the referral process. *Ergonomics*, 48, 1411-1422.

Greenley, J. R., Greenberg, J. S., & Brown (1997). Measuring quality of life a new and practical survey instrument. *Social Work,* 42(3), 244-254.

Grbich, C. (1999). *Qualitative research in health: An introduction.* Sydney: Allen and Unwin.

Guadagnoli, E., & Velicer, W. F. (1988). Relation of sample size to the stability of component patterns. *Psychological Bulletin*, 103, 265-275.

Guilford, J. P. (1946). New standards for test evaluation. *Educational and Psychological Measurement*, 6, 427-439.

Guilford, J. P. (1956) *Psychometric methods* New York: McGraw-Hill.

Guyatt, G. H, Feeny, D. H., & Patrick, D. L. (1993). Measuring health-related quality of life. *Basic Science Review* 118, 622-629.

Hague, P. (1993). *Questionnaire design*. London: Kogan Page Limited.

Hallal, P. C., & C. G. Victora (2004). Reliability and validity of the international physical activity questionnaire (IPAQ). *Medicine and Science in Sports and Exercise* 36, 556.

Hambleton, R. K., & Jones, R. W. (1993). Comparison of classical test theory and item response theory and their applications to test development. *Educational Measurement Issues and Practice*, 12(3), 38-47.

Hand, H. (2003). The mentor's tale: A reflexive account of semi-structured interviews. *Nurse Researcher*, 10(3), 15-27.

Hansen, E. C. (2001). Medical understandings of lifestyle: An interpretive study of lifestyle as a medical explanatory framework. Unpublished PhD Thesis. University of Tasmania.

Hansen, E. C. (2006). *Successful qualitative health research. A practical introduction.* London: Open University Press.

Hansson, L. (1999) Advancing the concepts of quality of life. In S. Priebe, J. P. L. Oliver, & W. Kaiser, (Eds.) *Quality of life and mental health care.* (pp.19-27). Petersfield and Philadelphia: Wrightson Biomedical Publishers.

Hardcastle, S. & Taylor, A. (2001). Looking for more than weight loss and fitness gain: Psychosocial dimensions among older women in a primary-care exercise referral program. *Journal of Aging and Physical Activity*, 9(3), 313-328.

Harman, H. H. (1976). *Modern factor analysis.* Chicago: University of Chicago Press.

Hattie, J. (1985). Methodology review: assessing unidimensionality of tests and items. *Applied Psychological Measurement,* 9, 139-164.

Haynes, M. (1999). Whose Life is it Anyway? *Women's Studies International Forum,* 22(6), 659-671.

Hayton, J. C., Allen, D. G., & Scarpello, V. (2004). Factor retention decisions in exploratory factor analysis: A tutorial in parallel analysis. *Organizational Research Methods*, 7, 191-205.

Heesch, K. C., Masse, L. C., & Dunn, A. L. (2006). Using rasch modeling to re-evaluate three scales related to physical activity: Enjoyment, perceived benefits and perceived barriers. *Health Education Research*, 21, 58-72.

Heinzelmann, P., & Bagley, R. W. (1970). Response to physical activity programmes and their effects on health behaviours. *Public Health Reports*, 85, 905-911.

Hilari, K., Lamping, D.L., & Smith, S. C. (2003). Stroke and aphasia quality of life scale-39 (SAQOL-39): Evaluation of acceptability, reliability and validity. *Stroke*, 34(8), 1944-1950.

Hilditch, J. R., Lewis, J., Peter, A., Maris, B. V., Ross, A., & Franssen, E…Dunn, E. (1996). A menopause-specific quality of life questionnaire: Development and psychometric properties. *Maturitas Journal of the Climacteric and Post Menopause*, 24, 161-175.

Hillsdon, M., Foster, C., Naidoo, B., & Crombie, H., (2004). *The effectiveness of public health interventions for increasing physical activity among adults: A Review of Reviews*. London: Health Development Agency.

Hilton, C. E. (2010). Physical activity for health: Adult recommendations, interventions and evaluation. In H. Blake (ed.). *Physical activity in rehabilitation and recovery* (p. 5-16). NovaScience: New York.

Hilton, C. E., Milton, K., & Bull, F. C. (2009). Let's get moving: A feasibility trial of a 'physical activity care pathway' in primary care settings [Abstract]. *Journal of Sport and Exercise Sciences.* 26(Suppl. 2) 43.

Hilton, C. E., & Poulter, E. (2009). Motivational interviewing and physical activity. *SportEX Health,* 19, 10-12.

Hobart, J., Lamping, D., Fitzpatrick, R., Riazi, A., & Thompson, A. (2001). The Multiple Sclerosis Impact Scale (MSIS-29): A new patient-based outcome measure. *Brain,* 124, 962–973.

Holloway, I. (1997). *Basic Concepts of qualitative research.* London: Blackwell Science.

Holmes, S. (2005). Assessing the quality of life – reality or impossible dream? A discussion paper. *International Journal of Nursing Studies,* 42(4), 493-501.

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika,* 30, 179-185.

Horsburgh, D. (2003). Evaluation of qualitative research. *Journal of Clinical Nursing,* 12, 307-312.

Hojat, M., Samuel, S., & Thompson, T. L. (1995). Searching for the lost key under the light of biomedicine: A triangular biopsychosocial paradigm may cast additional

light on medical education, research and patient care. In S. K. Majumdar, L. M. Rosenfield, D. B. Nash, & A. M. Audet (Eds.), *Medicine and Healthcare into the 21st Century* (pp. 310-325). Easton, PA: Pennsylvania Academy of Science.

Hunt, S. D., Sparkman, R. D., & Wilcox, J. B. (1982). The pretest in survey research: Issues and preliminary findings. *Journal of Marketing Research* 19, 269-273.

Hunt, S. M. (1997). The problem of quality of life. *Quality of Life Research*, 6, 205-212.

Hunt, S. M., McKenna, S. P., McEwen, J., Williams, J., & Papp, E. (1981).The Nottingham health profile: Subjective health status and medical consultations. *Social Science and Medicine*, 15, 221-229.

Hurley, A. E., Scandura, T. A., Schriesheim, C. A., Brannick, M. T., Seers, A., Vandenberg, R. J. & Williams, L. J. (1997). Exploratory and confirmatory factor analysis: Guidelines, issues, and alternatives. *Journal of Organizational Behavior*, 18, 667-683.

Isaacs, A. J., Critchley, J. A, Tai, S. S., Buckingham, K., Westley, D., & Harridge, S. D. & Gottlieb, J. M. (2007). Exercise evaluation randomised trial (EXERT): A randomised trial comparing GP referral for leisure centre-based exercise, community based walking and advice only. *Health Technology Assessment*, 11, 1-99.

Jadad, A. R. & O'Grady, L. (2008). How should health be defined? *British Medical Journal*, 337, a2900.

James, D.V.B., Johnston, L. H., Crone, D., Sidford, A. H., Gidlow, C., & Morris, C. & Foster, C. (2008). Factors associated with physical activity referral uptake and participation. *Journal of Sports Sciences*, 26(2), 217-224.

Jancey, J. Lee, A. Howat, P. Clarke, A. Wang, K. & Shilton, T. Reducing attrition in physical activity programmes for older adults. *Journal of Physical Activity and Aging*, 15, 152-165.

Keane, T. M., Caddell, J. M., & Taylor, K. L. (1988). Mississippi scale for combat-related posttraumatic stress disorder: Three studies in reliability and validity. *Journal of Counseling and Clinical Psychology*, 56, 85-90.

Keating, K. M. (1989). Self selection: are we beating a dead horse? *Evaluation and Program Planning*, 12(2), 137-142.

Keyes, C. L. M. (2005). Mental illness and/or mental health? Investigating axioms of the complete state model of health. *Journal of Consulting and Clinical Psychology*, 73(3), 539 – 548.

Kim, J. O., & Mueller, C. W. (1978). *Introduction to Factor Analysis: What it is and how to do it.* Beverly Hills, CA: Sage.

Kirshner, B., & Guyatt, G. (1985). A methodological framework for assessing health indices. *Journal of Chronic Disease*, 38, 1, 27-36.

Kitzinger, J. (1994). The methodology of focus group: The importance of interaction between research participants. *Sociology of Health and Illness* 16, 103–121.

Kitzinger, J. (2000). Focus groups with users and providers of health care. In C. Pope &, N. Mays (Eds.), *Qualitative research in healthcare.* London: British Medical Journal Books.

Kline, T. J. B. (2005). *Psychological testing: A practical approach to design and evaluation*. London: Sage.

Koch T. (1994). Establishing rigor in qualitative research: The decision trail. *Journal of Advanced Nursing* 19, 976–986.

Kolotkin, R. L., Crosby, R. D., Kosloski, K. D., & Williams, G. R. (2001). Development of a brief measure to assess quality of life in obesity. *Obesity Research*, 9(2), 102-111.

Krueger, R. A. (1993). Quality control in focus group research. In D. L. Morgan (Ed.), *Successful focus groups. Advancing the state of the art.* London: Sage.

Krueger, R. A. (1994). *Focus groups: A practical guide for applied research* (2nd ed.). London: Sage.

Kurtzke J. F. (1983). Rating neurologic impairment in multiple sclerosis: An expanded disability status scale (EDSS). *Neurology* 33, 1444–52.

Laitakari, J., & Asikainen (1998). How to promote physical activity through individual counseling – a proposal for a practical model of counseling on health-related physical activity. *Patient Education and Counseling,* 33(Supp 1.), 13 – 24.

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

Larson, R. (1978). Thirty years of research on the subjective well-being of older Americans. *Journal of Gerontology,* 33, 109-125.

Las Hayas, C. L., Quintana, J. M., Padierna, À., Bilbao, A., Munoz, P., & Madrazo…Cook, E.F. (2006). The New Questionnaire Health-Related Quality of Life for Eating Disorders Showed Good Reliability and Validity. *Journal of Clinical Epidemiology* 59, 192-200.

Lautenschlager, G. J. (1989). A comparison of alternatives to conducting Monte Carlo analyses for determining parallel analysis criteria. *Multivariate Behavioral Research,* 24, 365-395.

Lawton, B. A., Rose, S. B., Elley, C. R., Dowell, A. C., Fenton, A., & Moyes, S. A. (2008). Exercise on prescription for women aged 40-74 recruited through primary care: Two year randomised controlled trial. *British Medical Journal,* 337, 1-7.

Lee, A. S. W., Griffin, S. J., & Simmons, R. K. (2009). An evaluation of the effectiveness of 'active for life': An exercise referral scheme in West Suffolk. *Public Health,* 123, 670-672.

Lehman, A. F. (1983). The effects of psychiatric symptoms on quality of life assessments among the chronically mentally ill. *Evaluation and Programme Planning,* 6, 143-151.

Lehrner, J., Kalchmayr, R., Series, W., Olbrich, A., Pataraia, E., & Aull, S...Baumgartner, C. (1999). Health-related quality of life (HRQOL), activity of daily living (ADL) and depressive mood disorder in temporal lobe epilepsy patients. *Seizure, 8(2), 88-92.*

Leitz, P. (2010). Research into questionnaire design. A summary of the literature. *International Journal of Market Research* 52(2), 249-272.

Lejuez, C. W., Simmons, B. L., Aklin, W. M., Daughters, S. B. & Dvir, S. (2004). Risk-taking propensity and risky sexual behaviour of individuals in residential substance misuse treatment. *Addictive Behaviours.* 29, 1643-1647.

Le Masurier, G. C. & Tudor-Locke, C. (2003). Comparison of pedometer and accelerometer accuracy under controlled conditions. *Medicine and Science in Sports and Exercise*, 35(5), 867-871.

Lincoln, Y. S. & Guba E. (1985). *Naturalistic Enquiry.* London: Sage.

Liu, J. L. Y., Maniadakis, N., Gray, A., & Rayner, M. (2002). The economic burden of coronary heart disease in the UK. *Heart,* 88, 597-603.

Liu, O. L., & Rijmen, F. (2008). A modified procedure for parallel analysis of ordered categorical data. *Behavior Research Methods,* 40(2), 556-562.

Loftland, J., & Loftland, L. H. (1984). *Analysing social settings* (2nd ed.). Belmont: Wadsworth.

Lord, F. M. (1980). *Applications of item response theory to practical testing problems.* Mahwah, NJ: Erlbaum.

Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores.* Reading MA: Addison-Welsley Publishing Company.

Lozano, L. M., Garcia-Cueto, E., & Muniz, J. (2008). Effect of the number of response categories on the reliability and validity of rating scales. *Methodology*, 4(2), 73-79.

MacCallum, R. C., Widaman, K. F., Preacher, K. J., & Hong, S. (2001). Sample size in factor analysis: The role of model error. *Multivariate Behavioural Research*, 36, 611-637.

Marcus, B. H., & Forsyth, L. A. (2003). *Motivating people to become physically active.* Human Kinetics, Champaign, IL.

Marcus, B. H., Pinto, B. M., Simkin, L. R., Audrain, J. E., & Taylor, E. R. (1994). Application of theoretical models to exercise behaviour among employed women. *American Journal of Health Promotion*, 9, 49–55.

Marx, R. G. (2004). Comparing time intervals for test-retest reliability: Measuring stability and calculating a sample size. *Journal of Clinical Epidemiology* 57, 995.

Maslow, A. (1954). *Motivation and personality.* New York: Harper.

Maslow, A. H. (1968). *Toward a psychology of being* (2nd ed.). Princeton, NJ: Nostrand.

Mason, J. (1996). *Qualitative researching.* London: Sage.

Mattick, R. P., & Clarke, J. C. (1998). Development and validation of measures of social phobia scrutiny fear and social interaction anxiety. *Behaviour Research and Therapy*, 36(4), 455-470.

Mazurek Melnyk, B., & Fineout-Overholt, E. (2005). Creating a vision: motivating change to evidence based practice in individuals and organizations. In B. Mazurek Melnyk, and, E. Fiineout-Overholt (Eds.), *Evidence based health policy: Problems and possibilities.* London: Oxford University Press.

McAuley, E., & Courneya, K. (1994). The subjective exercise experiences scale (SEES): Development and preliminary validation. *Journal of Sport & Exercise Psychology,* 16, 163-177.

McDonald, R. P. (1985). *Factor Analysis and Related Methods.* Hillsdale, NJ: Lawrence Erlbaum Associates.

McKay, H. A., Macdonald, H., Reed, K. E., & Khan, K. M. (2003). Exercise interventions for health: Time to focus on dimensions, delivery and dollars. *British Journal of Sports Medicine*, 37, 98–99.

McKenna, S. P, Doward, L. C., Whalley, D., Tennant, A., Emery, P., & Veale, D. J. (2004). Development of the PsAQoL: A quality of life instrument specific to psoriatic arthritis. *Annals of the Rheumatic Diseases*, 63, 162-169.

Mcnair, D. M., Lorr, M. & Droppleman L. F. (1992) *Manual for the Profile of Mood States.* San Diego, California: Educational and Industrial Testing Service.

Miller, N. E. & Dollard, J. (1941). *Social learning and imitation*. New Haven: Yale University Press.

Miles, M. & Huberman, A. (1984). *Qualitative data analysis.* London: Sage.
Miller, W. R., & Rollnick, S. (2002). Motivational interviewing: Preparing people for change. London: Guilford Press.

Milne, H. M., Guilfoyle, A., Gordon, S., Wallman, K. E., & Courneya, K. S. (2007). Personal accounts of exercise and quality of life from the perspective of breast cancer survivors. *Quality of Life Research* 16, 1473-1481.

Mischler, E. G. (1986). *Research interviewing: Context and narrative.* London: Harvard University Press.

Mitchell, S. G., Peterson, J. A. & Kaya, S. (2004). Making the switch to digital audio. *International Journal of Qualitative Methods, 3, 4*. Retrieved August 23, 2010 from http://www.ualberta.ca/~iiqm/backissues/3_4/pdf/mitchell.pdf.

Montazeri, A., Goshtasebi, A. Vahdaninia, M., & Gandek, B. (2005). The short form health survey (SF-36): Translation and validation study of the Iranian version. *Quality of Life Research,* 14(3), 875-882.

Morgan, D. L. (1997). *Focus Groups as qualitative research* (2nd ed.). London: Sage.

Morgan, O. (2004). Approaches to increase physical activity: Reviewing the evidence for exercise referral schemes. *Journal of the Royal Institute for Public Health,* 119, 361-370.

Morse, J. M. (1989). *Qualitative nursing research: A contemporary Dialogue.* London: Sage.

Morten, M. (2010). Reducing socially desirable responses in epidemiologic surveys: An extension of the randomized response technique. *Epidemiology,* 21(3), 379-382.

Motl, R.W., & Snook, E. M. (2008). Physical activity, self-efficacy and quality of life in multiple sclerosis. *Annals of Behavioural Medicine* 35, 111-1155.

Muhr, T. (2000). Increasing the reusability of qualitative data with XML, *Forum Qualitative Research 1, 3.* Retrieved August 13, 2010 from http://www.qualitative-research.net/index.php/fqs/article/view/1037/2244.

Murphy, K. R., & Davidshofer, C. O., (2001). *Psychological testing: Principles and Applications*. Upper Saddle River: Pearson Prentice Hall.

Mutrie, N. (2000). The Relationship between physical activity and clinically defined depression. In S. J. H. Biddle, K. R. Fox, &, S. H. Boutcher (Eds.), *Physical Activity and Psychological Well-being* (pp. 46–62). London: Routledge.

Myers, A. M., Malott, O. W., Gray, E., Tudor-Locke, C., Ecclestone, N.A., & Cousins, S.O. et al. (1999). Measuring accumulated health-related benefits of exercise participation for older adults: The vitality plus scale. *Journal of Gerontology,* 54(9), 456-66.

Mykletun, A., Stordal, E., & Dahl, A. A. (2001). Hospital Anxiety and Depression (HAD) scale: factor structure, item analyses and internal consistency in a large population. *The British Journal of Psychiatry*, 179, 540, 544.

National Institute for Health and Clinical Excellence (2006a). *Four Commonly Used Methods to Increase Physical Activity: Brief Interventions in Primary Care, Exercise Referral Schemes, Pedometers, and Community-Based Exercise Programmes for Walking and Cycling.* London: The Stationery Office.

National Institute for Health and Clinical Excellence (2006b). *Implementation Advice: Four Commonly used Methods to Increase Physical Activity.* London: The Stationery Office.

National Institute for Health and Clinical Excellence (2006c). *A Rapid Review of the Effectiveness of Exercise Referral Schemes to Promote Physical Activity in Adults.* London: The Stationery Office.

National Institute for Health and Clinical Excellence (2007). *Behaviour change at population, community and individual levels.* London: The Stationery Office.

National Obesity Observatory (2009). *Standard evaluation framework for weight management interventions.* London: The Stationery Office.

Nijsten, T., Unaeze, J., & Stern, R. S. (2006). Refinement and reduction of the impact of psoriasis questionnaire: Classical test theory vs. rasch analysis. *Epidemiology and Health Services Research*, 154, 692-700.

Novick, M. R. (1966) The axioms and principal results of classical test theory. *Journal of Mathematical Psychology*, 3(1), 1-18.

Nuckols, R. C. (1953). A note on pretesting public opinion questions. *Journal of Applied Psychology*, 37, 119-120.

Nunnally, J. C. (1978). *Psychometric theory.* New York: Mc-Graw-Hill.

Nunnally, J. C., & Bernstein, I. H.(1994). *Psychometric theory* (3rd ed.). New York: McGraw- Hill.

Nupponen, R. (1998). What is counseling all about – basics in the counseling of health-related physical activity. *Patient Education and Counselling*, 33(Suppl.1) 61-67.

Oakley, A. (2001). Evaluating health promotion: methodological diversity. In S. Oliver, &, G. Peersman (Eds.) *Using Research for effective Health Promotion*, Buckingham: Open University Press.

O'Boyle, C. A. (1997). Measuring the quality of later life. *Philosophy Transactions of the Royal Society of London,* 352, 1871-1879.

O'Brien, S. F., Ram, S. S., Vamvakas, E. C., & Goldman, M. (2007). The Canadian blood donor health assessment questionnaire: Lessons from history, application of cognitive science principles, and recommendations for change. *Transfusion Medicine Reviews,* 21(3), 205-222.

Ohira, T., Schmitz, K. H., Ahmed, R. L., & Yee, D. (2006). Effects of weight training on quality of life in recent breast cancer survivors. The weight training for breast cancer survivors (WTBS) Study. *American Cancer Society* 106(9), 2076-2083.

Okochi, J., Utsunomiya, S., & Takahashi, T. (2005). Health measurement using the ICF: Test-retest reliability study of ICF codes and qualifiers in geriatric care. *Health and Quality of Life Outcomes,* 46(3), 1-13.

Oppenheim, A. N. (1992). *Questionnaire design, interviewing and attitude measurement new edition.* Continuum: London.

Pais-Ribeiro, J. L. (2004). Quality of life is a primary end-point in clinical settings. *Clinical Nutrition,* 23, 121-130.

Parfitt, G., Markland, D., & Holmes, C. (1994). Response to physical exertion in active and inactive males and females. *Journal of Sport and Exercise Psychology,* 16, 178-186.

Patton, M. O. (1990). *Qualitative evaluation and research Methods* (2nd ed.). London: Sage.

Pett, M. A., Lackey, N. R., & Sullivan, J. J. (2003). Making sense of factor analysis: The use of factor analysis for instrument development in health care research. London: Sage.

Petajan, J. H., Gappmaier, E., White, A. T., Spencer, M. K., Mino, L., & Hicks, R. (1996). Impact of aerobic training on fitness and quality of life in multiple sclerosis. *Annals of Neurology*, 39(4), 432-441.

Pinto B. M., Frierson G. M., Rabin C., Trunzo J. J., & Marcus B. H. (2005). Home-based physical activity intervention for breast cancer patients. *Journal of Clinical Oncology*, 23, 3577–3587.

Popay, J., Rogers, A., & Williams, G. (1998) Rationale and standards for the systematic review of qualitative literature in health services research. *Qualitative Health Research* 8(3), 341–351.

Presser, S., Couper, M. P, Lessler, J. T., Martin, E., Martin, J., & Rothgeb, J. M. & Singer, E. (2004). Methods for testing and evaluating survey questions. *Public Opinion Quarterly* 68(1), 109–30.

Prieto, L. Alonso, J. & Lamarca, R. (2003). Classical test theory versus rasch analysis for suality of life questionnaire reduction. *Health and Quality of Life*

Prochaska, J. O, DiClemente, C.C., & Norcross, J. C. (1992). In search of how people change : Applictions to addictive behaviours. *American Psychologist*, 47, 1102-1114.

Rabin, B. A., Brownson, R. C., Kerner, J. F., & Glasgow, R.E. (2006). Methodologic challenges in disseminating evidence based interventions to promote physical activity. *American Journal of Preventative Medicine,* 31(suppl. 4), 24-34.

Radloff, L. S. (1977) The CES-D scale: A self report depression scale for research in the general population. *Applied Psychological Measurement* 1, 385-401.

Ramp, M., Khan, F., Misajon, R. A., & Pallant, J. F. (2009). Rasch analysis of the multiple sclerosis impact scale (MSIS-29). *Health and Quality of Life Outcomes,* 7(58), 1-10.

Rejeski, W. J.,Brawley, L. R., & Shumaker, S. A. (1996). Physical activity and health-related quality of life. *Exercise and Sport Sciences Reviews,* 24, 71-108.

Rice, P. L. & Ezzy, D. (1999). *Qualitative research methods: A health focus.* London: Oxford University Press.

Richards, A., Kettelmann, K. K., and Ren, C. R. (2006). Motivating 18 to 24-Year-Olds to Increase their Fruit and Vegetable Consumption. *Journal of the American Diatetic Association,* 106, 1405-1411.

Riddoch, C., Puig-Ribera, A., & Cooper, A. (1998). *Effectiveness of physical activity schemes in primary care: A review*. London: Health Education Authority.

Riessman, C. K. (1993). *Narrative analysis.* London: Sage.

Robison, J. & Rogers, M. (1994). Adherence to exercise programmes. *Sports Medicine,* 17, 39–52.

Rodebaugh, T. L., Woods, C. M., & Heimberg, R. G. (2007). The reverse of social anxiety is not always the opposite: The reverse-scored items of the social interaction anxiety scale do not belong. *Behaviour Therapy*, 38(2), 192-206.

Rollnick, S. Miller, W. R., & Butler, C. C. (2008*). Motivational interviewing in health care: Helping patients change behaviour.* London: Guilford Press.

Romberg, A., Virtanen, A., & Ruutiainen, J. (2005). Long-term exercise improves functional impairment but not quality of life in multiple sclerosis. *Journal of Neurology*, 252(7), 839-845.

Roy D. J (1992). Measurement in the service of compassion. *Journal of Palliative Care* 8(3), 4-51.

Rubin, H. J., Rubin, I. S. (2004). *Qualitative interviewing the art of hearing data.* (2nd ed.). London: Sage.

Rust, J., &, Golombok, S. (2009). *Modern psychometrics: The science of psychological Assessment*: London: Routledge.

Saracci, R. (1997). The World Health Organisation needs to reconsider its definition of health. *British Medical Journal*, 314, 1409-1410.

Schag C. A., Ganz P. A., & Heinrich R. L. (1991). Cancer rehabilitation evaluation system - short form (CARES-SF). A cancer specific rehabilitation and quality of life instrument. *Cancer*, 68, 1406–1413.

Schunemann, H. J., Sperati, F., Barba, M., Santesso, N., Melegari, C., & Akl, E. A…Muti, P. (2010). An instrument to assess quality of life in relation to nutrition: Item generation, item reduction and initial validation. *Health and Quality of Life Outcomes*, 8(26), 2-13.

Schwandt, T. A. (2001). *Dictionary of qualitative inquiry* (2nd ed.). London: Sage.

Schwartz, C. E., Coulthard-Morris, L., Zeng, Q., & Retzlaff, P. (1996). Measuring self-efficacy in people with multiple sclerosis: A validation study. *Archives of Physical Medicine and Rehabilitation.* 77, 394–398.

Shacklock, G., & Smyth, J. (1998). *Being Reflexive in Critical Educational and Social Research.* London: Taylor and Francis.

Silverman, D. (2000). *Doing qualitative research. A practical handbook.* London: Sage.

Singh, S. (1997). Why are GP exercise schemes so successful (for those who Attend)? Results from a pilot study. *Journal of Management in Medicine*, 11(4), 233-237.

Sijtsma, K. & van der Ark, L. (2003). Investigation and treatment of missing item scores in test and questionnaire data. *Multivariate Behavioural Research*, 38(4), 505-528.

Skevington, S. M. (1999). Measuring quality of life in Britain: Introducing the WHOQOL-100. *Journal of Psychosomatic Research* 47, 5449-5459.

Skevingon, S. M., Lofty, M., & O'Connell, K.A. (1994). The World Health Organization's WHOQOL-BREF quality of life assessment: Psychometric properties and results of the international field trial. A report from the WHOQOL Group. *Quality of Life Research,* 13(2), 299-310.

Skevington, S. M., & Tucker, C. (1999). Designing response scales for cross-cultural use in health care: Data from the development of the UK WHOQOL. *British Journal of Medical Psychology*, 72, 51-61.

Skinner, B. F. (1938). *The behaviour of organisms*. New York: Appleton-Century-Crofts.

Snaith, P. R. (2003). The hospital anxiety and depression scale. *Health and Quality of Life Outcomes,* 1(29), 1-4.

Spall, S., & Stephen, F. (1998). Peer debriefing in qualitative research: Emerging operational models, *Qualitative Enquiry*, 4(2), 280-292.

Spies, J. B., Coyne, K., Guau, N. G., Boyle, D., Skyrnarz-Murphy, K., & Gonzalves, S.M. (2002). The UFS-QOL, a new disease-specific symptom and health-related quality of life questionnaire for leiomyomata. *Obstetrics and Gynecology* 99(2), 290-300.

Springett, J., & Dugdill, L., (1999). Health promotion programmes and policies in the workplace: A new challenge for evaluation. *European Health Promotion Series no. 7, Copenhagen: Health Promotion Wales/World Health Organization.*

Stathi, A., Fox, K. R., & McKenna, J. (2002). Physical activity and dimensions of subjective well-being in older adults. *Journal of Aging and Physical Activity,* 10, 76-92.

Stathi, A. McKenna, J. & Fox, K. R. (2003). The experiences of older people participating in exercise referral schemes. *The Journal of the Royal Society for the Promotion of Health*, 124(1), 18-23.

Stevens, J. P. (2002). *Applied multivariate statistics for the social sciences* (4th ed.). Hillsdale NJ: Erlbaum.

Stewart, K. J., Turner, K. L., Bacher, A. C., DeRegis, J. R., Sung, J., & Tayback, M. & Ouyang, P. (2003). Are fitness, activity, and fatness associated with health-related quality of life and mood in older persons? *Journal of Cardiopulmonary Rehabilitation,* 23(2), 115-121.

Stewart, S. T., Cutler, D. M., & Rosen, A.B. (2010). Forecasting the effects of obesity and smoking on U.S. life expectancy. *The New England Journal of Medicine,* 4, 855-857.

Stineman, M. G., Shea, J. A., Jette, A., Tassoni, C. J., Ottenbacher, K. J., & Fiedler, R. (1996) The functional independence measure: Tests of scaling assumptions, structure, and reliability across 20 diverse impairment categories *Archives of Physical Medicine and Rehabilitation* 77, 1101-1108

Streiner, D. L., & Norman, G. R. (2008). *Health measurement scales: A practical guide to their development and use.* (4th ed.). New York: Oxford University Press.

Taylor, A. H. (2000). Physical activity, anxiety and stress. In S. J. H. Biddle, K. R. Fox, & S. H. Boutcher (Eds.), *Physical Activity and Psychological Well-being* (pp. 10–45). London: Routledge.

Taylor, A. H. (2003). The role of primary care in promoting physical activity. In McKenna, J., & Riddoch, C. (Eds.), *Perspectives on Health and Exercise* (pp. 153-173). Basingstoke: Palgrave Mcmillan.

Taylor, A. H., Doust, J., & Webborn, A.D.J. (1998). Randomised control trial to examine the effects of a GP exercise referral programme in East Sussex, UK, on modifiable coronary heart disease risk factors. *Journal of Epidemiology and Community Health,* 52, 595-601.

The WHOQOL Group (1998). The World Health Organisation quality of life assessment (WHOQOL): Development and general psychometric properties. *Social Science and Medicine,* 46(12), 1569-1585.

Thomas, D. R. (2006). A general inductive approach for analyzing qualitative evaluation data. *American Journal of Evaluation* 27(2), 237-246.

Thomsen, T. F., Davidsen, M. Ibsen, H., Jørgensen, T., Jensen, G., & Borch-Johnsen, K. (2001). A new method for CHD prediction and prevention based on regional risk scores and randomized clinical trials; PRECARD(R) and the Copenhagen risk score. *Journal of Cardiovascular Risk* 8(5), 291-297.

Thompson, B. (1995). Ethical dimensions in trauma research. *American Sociologist,* 26(2), 54-69.

Trigg, R., Jones, R. W., & Skevington, S. M. (2007). Can people with mild to moderate dementia provide reliable answers about their quality of life? *Age and Ageing,* 36, 663-669.

Vickrey, B. G., Hays, R. D., Harooni, R., Myers, L. W., & Ellison, G. W. (1995). A health-related quality of life measure for multiple sclerosis. *Quality of Life Research,* 4, 187-206.

Vogt, W. P. (1993). *Dictionary of statistics and methodology: A nontechnical guide for the social sciences.* Newbury Park, CA: Sage.

Warburton, D. E. R., Nicol, C. W., & Bredin, S. S. D. (2006). Health benefits of physical activity: The evidence. *Canadian Medical Journal Association*, 174, 801-809.

Ware, J. E. J. (1996). The SF-36 health survey. In Spiker, B. *(*Ed.), *Quality of Life and Pharmacoeconomics in Clinical Trials* (2nd ed.). (p. 337-345). Philadelphia, PA: Lippincot-Raven.

Ware, J. E. J. Kosinski, M., & Keller, S. D. (1996). A 12-Item short form health survey. *Medical Care* 34(3), 220-223.

Warnecke, R. B., Johnson, T. P., Chavez, N., Sudman, S., O'Rourke, D.P., & Lacey L. & Horm, J. (1997). Improving question wording in surveys of culturally diverse populations. *Annals of Epidemiology,* 7(5), 334-342.

Webb, C. (1992). The use of the first person in academic writing: Objectivity, language and gatekeeping. *Journal of Advanced Nursing* 17, 747–752.

Webb, C. (1996). Caring, curing, coping: Towards an integrated model. *Journal of Advanced Nursing* 23, 960–968.

Wee, C. C., Davis, R. B., & Hamel, M. B. (2008). Comparing the SF-12 and SF-36 health status questionnaires in patients with and without obesity. *Health and Quality of Life Outcomes,* 6(11), 1-7.

Weitzman, E. A. (1999). Analyzing qualitative data with computer software. *Health Services Research,* 34(5), 166-178.

Wildy, H., & Clarke, S. (2009). Using cognitive interviews to pilot an international survey of principal preparation: A Western Australian Perspective. *Educational Assessment, Evaluation and Accountability,* 21, 105-117.

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.

Willis, G. B. (2005). *Cognitive interviewing. A tool for improving questionnaire design.* London: Sage.

Wilkinson, S. (2003). Focus groups. In J.A. Smith (Ed.), *Qualitative psychology: A practical guide to research methods.* (pp.184-204). London: Sage.

Wilson, I. B. & Cleary, P. D. (1995). Linking clinical variables with health-related quality of life. A conceptual model of patient outcomes. *The Journal of the American Medical Association,* 273(1), 59-65.

Wilson, M., Allen, D. D., & Li, J. C. (2006). Improving measurement in health hducation and health behaviour research using item response modeling: Comparison with the classical test theory approach. *Health Education Research,* 21, 19-32.

Winett, R. A. (1998). Developing more effective health-behavior programs: analyzing the epidemiological and biological bases for activity and exercise programs. *Applied and Preventive Psychology,* 7, 209-224.

Woods, M. D., Agarwal, S. Young, B. Jones, D. & Sutton, A. (2004). Integrative approaches to qualitative and quantitative evidence. London: Health Development Agency.

World Health Organisation. (1946). *Constitution of the World Health Organisation.* Geneva: World Health Organization.

World Health Organization (2001). *International classification of functioning, disability and health*. Geneva: World Health Organization.

Yaremko, R. M., Harari, H., Harrison, R. C., & Lynn, E. (1986). *Handbook of research and quantitative methods in psychology: For students and professionals.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Yesavage, J. A., Brink, T. L., Rose, T. L., Lum, O., Huang, V., & Adey, M. & Leirer, V. O. (1982). Development and validation of a geriatric depression screening scale: A preliminary report. *Journal of Psychiatric Research*, 17(1), 37-49.
Outcomes, 1(27), 1-13.

Appendices

Appendix 1 NRES Participant Information Sheet: Focus Groups

Part 1

You are being invited to take part in a research study. Before you decide it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully. Talk to others about the study if you wish.

The number of exercise referral programmes just like the scheme you have been referred into are growing. To date, there is no standardised method of evaluating these programmes which is essential for their continuation. It is also important that there is adequate opportunity for the participants to explore their experiences in order that the reasons for continuing to be physically active for health may be understood and individuals have an appropriate opportunity to explore their personal benefits and barriers to a more active lifestyle

This research study aims to explore the experiences of individuals who are referred into a programme of exercise via the use of group interviews (focus groups). The reports made by these participants will be used to formulate a questionnaire which will be used as an essential tool for exercise referral programme evaluation. You are being requested to contribute to a focus group.

Please ask your instructor, the exercise promotion officer or the primary investigator if you would like more information. Take time to decide whether or not you wish to take part.

What is the purpose of the study?

This research is primarily for educational purposes. The primary researcher is undertaking a PhD at Nottingham Trent University. It is expected that the outcome of this research will also help Positive Moves and other referral schemes assess the needs of their clients and evaluate their programmes more effectively.

Why have I been chosen?

You have been chosen to attend a focus group because you have completed the initial prescription period of your referral into exercise and you are in an optimum position to convey your thoughts and feelings regarding your experiences of increasing your physical activity levels.

Do I have to take part?

No. It is up to you to decide whether or not to take part. If you do, you will be given this information sheet to keep and be asked to sign a consent form. You are still free to withdraw at any time and without giving a reason. A

decision to withdraw at any time, or a decision not to take part, will not effect the standard of care you receive.

What will happen to me if I take part?

- The focus group meeting will be arranged at a Gedling Leisure Centre, locally.
- In total, the group meeting will last for approximately 2 hours; including refreshments, consent form completion, the focus group itself and an opportunity to speak with the primary investigator before and after the group discussion should you wish.
- It is expected that the actual group discussion will last for approximately 1-hour (this is included in the total 2-hour period).
- You will be invited to sit with the other 3-9 men and women participating in the group. The primary investigator will prompt discussion using open ended questions intended to encourage the group to comment on their health related, physical, emotional, practical, social and environmental experiences of exercise referral, for example.
- 4-6 focus groups will be arranged in total. You will only attend one of these.
- The primary researcher will use the recordings of these group discussions to identify common themes expressed by participants. These themes will be used to formulate a questionnaire which is intended to contribute to a key element of the evaluation of exercise referral schemes for participants across the UK.
- You may be contacted by the primary investigator during the transcription phase of the study in order to establish that what you said has been interpreted as you meant it.
- Some referral schemes use focus groups to identify the strengths and weaknesses of their scheme. However, this study will use the data produced from the focus groups to formulate a standard tool used for evaluative purposes.
- The audio, visual and written recordings taken from the focus group will be subject to the strictest of confidentiality. Written and audio material will be coded to ensure anonymity; this means that your name will never be used in any written material. All data, audio, visual and written will be placed in a locked filing cabinet to which only the primary investigator will have access.
- If the results of the research are shared at academic conferences, journals and or public media, you will not be identified by name or otherwise unless consent is given by you specifically following the completion of the research.
- There is no long-term monitoring or follow up planned specifically. However, if you wish to contact the research team during the 3-year study period, you are welcome to do so.

What do I have to do?

If you decide to attend the focus group, after completing a consent form and meeting the rest of the participants, you will be invited to join the discussion. You will be asked to sit with the other members of the group and the primary investigator will prompt discussion by asking the group, questions related to

your experience of exercise referral. You will not be expected to evaluate the scheme as such, although if you feel that the design of the scheme was an important factor which contributed to your responses to exercise referral, an exploration of this will be encouraged by the primary investigator. The overall purpose is to explore your experiences and the questions you will be asked will reflect this. You will not be expected to disclose any information you do not wish to and the questions will be designed to allow for your decision as to the level of personal information you choose to disclose.

What are the possible benefits of taking part?

There is no direct clinical benefit from your participation. However, your participation is likely to assist the future development and evaluation procedures of exercise referral schemes and the experiences of those referred into them.

What happens when the research study stops?

When the research stops, the results of the research may be published in academic journals, local authority and Primary Care Trust publications and communicated at appropriate conferences. If you wish to know more about the course of the research, you may contact the primary investigator during the proposed 3-year period (the contact details of whom are given in part 2)

What if there is a problem?

Any complaint about the way you have been dealt with during the study or any possible harm you might suffer will be addressed. The detailed information regarding this is given in part two where you will also find a contact number for complaints.

Will my taking part in the study be kept confidential?

Yes. All the information regarding your participation in this study will be kept confidential. The details are included in part two.

Contact details

The primary investigator of this research is Charlotte Hilton. You may contact her during the hours of 9.30am and 5.30pm.

Address:
Charlotte Hilton
PhD Research Student
School of Biomedical and Natural Sciences
Nottingham Trent University
Nottingham
NG11 8NS
Telephone: 0115 8486601
Email: charlotte.hilton@ntu.ac.uk

Part Two

What will happen if I don't want to carry on with the study?

If you withdraw from the study, we will destroy all your identifiable data, however, we will need to use the data collected up until your withdrawal.

What if there is a problem?

Should you experience a problem either during or following your participation in the focus group, you may contact the primary investigator. Alternatively, you may wish to contact a member of the research team:

Primary Research Supervisor:
Dr. Antoinette Minniti
School of Biomedical and Natural Sciences
Nottingham Trent University
Nottingham
NG11 8NS
Telephone: 0115 8483918
Secondary Research Supervisor:
Professor Mark G. Darlison
Professor of Neuroscience
School of Biomedical and Natural Sciences
Nottingham Trent University
Nottingham
NG11 8NS
Telephone: 0115 8483207

Secondary Research Supervisor:
Dr. Stuart Mireylees
Principal Lecturer
School of Biomedical and Natural Sciences
Nottingham Trent University
Nottingham
NG11 8NS
Telephone: 0115 8483189

Complaints

If you have a concern about any aspect of this study, you should ask to speak with the researchers (contact details above) who will do their best to answer your questions. If you remain unhappy and wish to complain formally, you can do this through the Nottingham Trent University complaints procedure. Details can be obtained from the university.

Harm

In the event that something goes wrong and you are harmed during the research study there are no special compensation arrangements. If you are

harmed and this is due to someone's negligence then you may have grounds for a legal action for compensation against Nottingham Trent University but you may have to pay your legal costs. The normal National Health Service complaints mechanisms will still be available to you.

Will my taking part in this study be kept confidential?

Yes. The data for this study will be collected by transcribing word for word what is said during the focus groups. Transcription will be performed by the primary investigator and a professional transcription service. This process will be overseen by the primary investigators research supervisor. These are the only individuals who will have access to this data before it is coded (made anonymous) and also thereafter. Once coded, the full research team (identified above) may have access to the data. At this stage you will not be identified by the text generated from the transcription alone. In order to fully understand what is meant by the statements you make during the focus group, it may be necessary to listen to the audio tapes and view the visual data, this is the only way you will be identified. However, this material will be reviewed under strict confidential circumstances. Only the primary investigator and their research supervisor will have access to the audio recordings.

The custodian for the data is the primary investigator and their research supervisor (Dr. Antoinette Minniti). The data will be stored in locked filing cabinets to which only the primary investigator and their research supervisor will have access. The data will not be retained for use in future studies. The data will be retained for up to 3-years and disposed of securely.

What will happen to the results of the research study?

It is intended to communicate the results of the study overall within academic journals and conferences. There may also be an opportunity to communicate the findings within Gedling Borough Council's 'Contacts' magazine and Primary Care Trust (PCT) publications. You will not be identified in any presentation or publication unless you have consented to release such information.

Who is organising and funding the research?

Nottingham Trent University is part-funding the research by paying the academic fees of the primary researcher. There is a small stationery budget and all other costs are met by the primary researcher.

Who has reviewed the study?

This study was given a favourable ethical opinion for conduct in the NHS by an NHS Research Ethics Committee (REC) and Nottingham Trent University ethics.

You will be given a further copy of this information sheet and a signed consent form to keep before commencement of the focus group should you

choose to participate. Thank you for considering to take part and taking the time to read this information.

NOTTINGHAM TRENT UNIVERSITY

Appendix 2 Information for Participants : Focus Group

You have been asked to participate in this research because you have been referred for exercise to improve your health.

The complete research study aims to develop a questionnaire to help us understand aspects of quality of life that may be relevant or important to someone undertaking an exercise programme.

If you decide to take part, you will be asked to discuss your experiences of the exercise programme to which you were referred. The focus group facilitator (Charlotte) will use an interview guide to help prompt discussion. The focus group will be recorded. However you will not be identified in any of the materials that is used for this research or any publications.

If you have any questions once the session is completed, you may speak with Charlotte or contact:

Charlotte Hilton:
Email: charlotte.hilton@ntu.ac.uk
Telephone: 07939247921

Toni Minniti:
Email: antoinette.minniti@ntu.ac.uk
Telephone: 0115 88483918

Appendix 3 Consent Form

Participant Name:

CONSENT FORM

Name of Primary Researcher: Charlotte Hilton

1. I confirm that I have read and understand the information sheet for the study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily. ☐

2. I understand that I am providing consent for the use of audio taping, with possible use of verbatim quotation. ☐

3. I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason, without my medical care or legal rights being affected. ☐

4. I agree to take part in the above study ☐

Name of Participant	Date	Signature
Primary Investigator	Date	Signature

When completed, 1 for participant, 1 for primary investigator.

Appendix 4 NVivo Node Report

NVivo revision 2.0.163 Licensee: Information Systems

Project: Charlotte PhD 30th June 08 User: Administrator Date: 01/09/2009 - 14:59:34

NODE LISTING

Nodes in Set: All Nodes

Created: 30/06/2008 - 15:44:54

Modified: 30/06/2008 - 15:44:54

Number of Nodes: 70

1 Achievement

2 Affordability

3 Assistance with Surgery

4 Changes in Energy Levels

5 Changes in Lifestyle

6 Changes in Positive Mental Outlook

7 Choice of Activity

8 Choice of Time to Exercise

9 Comfortable with Self

10 Comparison to Others

11 Competitiveness/Determination

12 Healthy Eating Habits

13 Enjoyment

14 ER to Initiate PA

15 Expectations of Gym Environment

16 Fear of Exercise

17 Fear of Gym

18 Fear of Injury

19 Feeling Better

20 Feeling Better about Self

21 Feeling Healthier (Mentally and Physically)

22 Feeling Younger

23 Good Instructors
24 Disease Management
25 Helpful Leisure Centre Staff
26 History of PA
27 Importance of Social Support
28 Importance of Time for Self
29 Improved Sleep
30 Improvements in Active Daily Living
31 Improvements in Confidence and Self-Efficacy
32 Improvements in Independence
33 Improvements in Memory
34 Improvements in Motivation
35 Improvements in Physical Health
36 Improvements in QoL
37 Improvements in Self-Esteem
38 Improvements in Social Contact
39 Improvements in Stress Management
40 Improvements in Well-Being
41 Knowledge of Illness
42 Knowledge of Limitations
43 Knowledge of PA
44 Maintaining Exercise Behaviour
45 Managing Depression
46 Opportunity for PA
47 PA for Holistic Health
48 Pain Management
49 Using Gym Equipment
51 (1 1) /Good Instructors/Importance of Knowledge
52 (1 2) /Good Instructors/Importance of Support
53 (1 3) /Good Instructors/Importance of Communication
54 (1 4) /Good Instructors/Importance of Familiarity
55 (2 1) /QoL/Happy Fit and Healthy
56 (2 2) /QoL/Fit to Function
57 (2 3) /QoL/Energy Level, Illness
58 (2 4) /QoL/Family Time
59 (2 5) /QoL/Increased Importance with Age

(2 6) /QoL/Absence of Pain
(2 7) /QoL/Health
(2 8) /QoL/Income
(3 1) /Weight/Loss
(3 2) /Weight/Gain
(3 3) /Weight/Maintenance
(4 1) /Potential Barriers to PA/Poor Health
(4 2) /Potential Barriers to PA/Time
(4 3) /Potential Barriers to PA/Childcare
(4 4) /Potential Barriers to PA/Transport
(4 5) /Potential Barriers to PA/Weather

Appendix 5 NRES Participant Information Sheet: Questionnaire Pre-Testing

Part 1

You are being invited to take part in a research study. Before you decide it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully. Talk to others about the study if you wish.

The number of exercise referral programmes just like the scheme you have been referred into are growing. To date, there is no standardised method of evaluating these programmes which is essential for their continuation. It is also important that there is adequate opportunity for the participants to explore their experiences in order that the reasons for continuing to be physically active for health may be understood and individuals have an appropriate opportunity to explore their personal benefits and barriers to a more active lifestyle

This research study aims to explore the experiences of individuals who are referred into a programme of exercise via the use of group interviews (focus groups). The reports made by these participants has been used to formulate a questionnaire which will be used as an essential tool for exercise referral programme evaluation. You are being requested to contribute to the development of the questionnaire.

Please ask your instructor or the primary investigator if you would like more information. Take time to decide whether or not you wish to take part.

What is the purpose of the study?

This research is primarily for educational purposes. The primary researcher is undertaking a PhD at Nottingham Trent University. It is expected that the outcome of this research will also help other referral schemes assess the needs of their clients and evaluate their programmes more effectively.

Why have I been chosen?

You have been chosen to complete the questionnaire because you have been referred for exercise and you are in an optimum position to convey your thoughts and feelings regarding how well the questionnaire under development captures what is important to people like yourself.

Do I have to take part?

No. It is up to you to decide whether or not to take part. If you do, you will be given this information sheet to keep and be asked to sign a consent form. You are still free to withdraw at any time and without giving a reason. A decision to withdraw at any time, or a decision not to take part, will not affect the standard of care you receive.

What will happen to me if I take part?

- You will be asked to complete a questionnaire for exercise referral schemes that is under development.
- You will be asked to think aloud while completing the questionnaire to help us understand how you arrived at the response you gave and if we have managed to word the question so that it is interpreted as it was intended.
- You will also be asked additional questions as you complete the questionnaire to further this understanding.
- The audio and written recordings taken from the session will be subject to the strictest of confidentiality. Written and audio material will be coded to ensure anonymity; this means that your name will never be used in any written material. All data (audio and written) will be placed in a locked filing cabinet to which only the primary investigator will have access.
- If the results of the research are shared at academic conferences, journals and or public media, you will not be identified by name or otherwise unless consent is given by you specifically following the completion of the research.
- There is no long-term monitoring or follow up planned specifically. However, if you wish to contact the research team during the 3-year study period, you are welcome to do so.

What do I have to do?

If you decide to complete the questionnaire, after completing a consent form you will be guided through the questions with the primary investigator for this research who will encourage you to think aloud while responding and also ask you additional questions to help further our understanding of the questionnaire design.

What are the possible benefits of taking part?

There is no direct clinical benefit from your participation. However, your participation is likely to assist the future development and evaluation procedures of exercise referral schemes and the experiences of those referred into them.

What happens when the research study stops?

When the research stops, the results of the research may be published in academic journals, local authority and Primary Care Trust publications and communicated at appropriate conferences. If you wish to know more about the course of the research, you may contact the primary investigator during the proposed 3-year period (the contact details of whom are given in part 2)

What if there is a problem?

Any complaint about the way you have been dealt with during the study or any possible harm you might suffer will be addressed. The detailed information regarding this is given in part two where you will also find a contact number for complaints.

Will my taking part in the study be kept confidential?

Yes. All the information regarding your participation in this study will be kept confidential. The details are included in part two.

Contact details

The primary investigator of this research is Charlotte Hilton. You may contact her during the hours of 9.30am and 5.30pm.

Address:
Charlotte Hilton
PhD Research Student
School of Science and Technology
Nottingham Trent University
Nottingham
NG11 8NS
Telephone: 07939247921
Email: charlotte.hilton@ntu.ac.uk

Part Two

What will happen if I don't want to carry on with the study?

If you withdraw from the study, we will destroy all your identifiable data, however, we will need to use the data collected up until your withdrawal.

What if there is a problem?

Should you experience a problem either during or following your participation in the focus group, you may contact the primary investigator. Alternatively, you may wish to contact a member of the research team:

Primary Research Supervisor:
Dr. Antoinette Minniti
School of Science and Technology
Nottingham Trent University
Nottingham
NG11 8NS
Telephone: 0115 8483918

Complaints

If you have a concern about any aspect of this study, you should ask to speak with the researchers (contact details above) who will do their best to answer your questions. If you remain unhappy and wish to complain formally, you can do this through the Nottingham Trent University complaints procedure. Details can be obtained from the university.

Harm

In the event that something goes wrong and you are harmed during the research study there are no special compensation arrangements. If you are harmed and this is due to someone's negligence then you may have grounds for a legal action for compensation against Nottingham Trent University but you may have to pay your legal costs. The normal National Health Service complaints mechanisms will still be available to you.

Will my taking part in this study be kept confidential?

Yes. The data for this study will be collected by reviewing the audio recordings of feedback made by respondents as they complete the questionnaire. You will not be identified by the text generated from the analysis of the audio data or any notes taken during your completion of the questionnaire. Only the primary investigator and their research supervisor will have access to the audio tape recordings.

The custodian for the data is the primary investigator and their research supervisor (Dr. Antoinette Minniti). The data will be stored in locked filing cabinets to which only the primary investigator and their research supervisor will have access. The data will not be retained for use in future studies. The data will be retained for up to 3-years and disposed of securely.

What will happen to the results of the research study?

It is intended to communicate the results of the study overall within academic journals and conferences. There may also be an opportunity to communicate the findings within Primary Care Trust (PCT) publications. You will not be identified in any presentation or publication unless you have consented to release such information.

Who is organising and funding the research?

Nottingham Trent University is part-funding the research by paying the academic fees of the primary researcher. There is a small stationery budget and all other costs are met by the primary researcher.

Who has reviewed the study?

This study was given a favourable ethical opinion for conduct in the NHS by an NHS Research Ethics Committee (REC) and Nottingham Trent University ethics.

Thank you for considering to take part and taking the time to read this information.

NOTTINGHAM TRENT UNIVERSITY

Appendix 6 Pre-Testing Information for Participants: Self Complete

You have been asked to participate in this research because you have been referred for exercise to improve your health.

The complete research study aims to develop a questionnaire to help us understand aspects of quality of life that may be relevant or important to someone undertaking an exercise programme.

If you decide to take part, you will be asked to complete the questionnaire under development. While completing the questionnaire you will be encouraged to think aloud to help us understand how you arrived at the response you gave and if we have managed to word the question so that it is interpreted as it was intended. You will also be asked additional questions as you complete the questionnaire to further this understanding.

The session will be recorded and the personal details you provide will remain confidential. It is expected that this complete process will take less than one hour.

If you have any questions once the session is completed, you may contact the primary researcher Charlotte Hilton:

Telephone: 0115 8486601

Email: charlotte.hilton@ntu.ac.uk

Appendix 7 Pre-Testing Information for Participants : Interview/Telephone

You have been asked to participate in this research because you have been referred for exercise to improve your health.

The complete research study aims to develop a questionnaire to help us understand aspects of quality of life that may be relevant or important to someone undertaking an exercise programme.

If you decide to take part, you will be asked to complete the questionnaire under development. The questionnaire will be read to you and you will be asked to respond to each of the questions you are asked. At various points you will be asked to advise on the clarity of the questions asked and how difficult the questions are to answer, for example. This will help us understand how you arrived at the response you gave and if we have managed to word the question so that it is interpreted as it was intended. You will also be asked additional questions as you respond to the questions you are asked to further this understanding.

The session will be recorded and the personal details you provide will remain confidential. It is expected that this complete process will take less than one hour.

If you have any questions once the session is completed, you may contact the primary researcher Charlotte Hilton:

Telephone: 0115 8486601

Email: charlotte.hilton@ntu.ac.uk

Appendix 8 Interviewer Probe Protocol 1

This questionnaire is being developed to help us understand aspects of quality of life that may be relevant or important to someone undertaking an exercise programme.

Please answer all the questions you are asked. If you are unsure about which response to give to a question, please choose the **ONE** that seems most appropriate. This can often be your first response. At some points during the questionnaire response options will change so we ask that you answer carefully.

We ask that you think about your life in the last two weeks when responding to the questions that will follow. Please **circle** one number only.

Before we move on to the actual questions, can you tell me what this introduction is telling you?

1) How important to you is the affordability of leisure facilities in order to participate in a structured exercise programme?	What does the term 'structured exercise mean to you?
2) How much do you feel you have a choice regarding the exercise you undertake?	General probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
3) How much do you feel you have a choice regarding the time at which you exercise?	General probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
4) How much do you feel that lifestyle factors (e.g., transport, time, childcare, poor health & the weather) affect your ability to be physically active?	How easy or hard was it to choose an answer?
5) How much do you feel you compare yourself to others while exercising?	General probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.

6) How much do you worry about exercising in an environment with others?	General probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
7) How much do you worry about participating in structured exercise?	What does the term "structured exercise" mean to you?
8) How much do you worry about the prospect of exercising in a gym?	General probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
9) How much does any injury you may have prevent you from being physically active?	General probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking. Check the response scale is suitable for those who do not have any injury.
10) How much do you currently enjoy physical activity?	General probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
11) How much is social contact with others a part of your current lifestyle?	General probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
12) How important is it for you to manage your weight?	General probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
13) How much do you feel that you incorporate physical activity into your daily lifestyle?	What does the term "physical activity" mean to you? In what way does this differ to the word "exercise", if at all?
14) How well do you feel you adhere to eating habits that are beneficial to your health and any illness you may have?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.

15) How competitive are you?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
16) How determined are you?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
17) How confident are you using gym equipment?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
18) How confident are you that you can exercise in a leisure centre with minimum support?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
19) How confident are you in your ability to participate in regular physical activity and exercise?	Can you repeat the question I just asked you in your own words?
20) How confident are you in your ability to undertake daily tasks that do not necessarily involve physical activity?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
21) In general, how confident are you around around other people?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
22) How confident are you in your own physical ability?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
23) How confident are you to join in activities with family and friends?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
24) How important is it to you right now to maintain a physically active lifestyle?	What does the term “physically active lifestyle” mean to you?

25) How confident are you, right now that you can maintain a physically active lifestyle?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
26) How motivated are you to maintain a physically active lifestyle?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
27) How likely are you to maintain a physically active lifestyle?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
28) In general, how much opportunity do you have to make time for yourself?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
29) How often do you experience physical pain?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
30) How often do you feel you have to manage any symptoms of pain?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
31) How often do you feel you have to manage the symptoms of any illness you have, if any?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
32) How often do you feel you have achieved something in terms of your physical fitness?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
33) How would you rate your energy levels?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.

34) How would you rate your outlook on life?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
35) How would you rate your level of self-esteem?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
36) How would you rate your physical health?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
37) How would you rate your mental health?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
38) How would you rate your sense of well-being?	What does the term "well-being" mean to you?
39) How would you rate the quality of support you receive from others to be physically active?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
40) How would you rate the quality of your sleep?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
41) How would you rate your memory function?	What does the term "memory function" mean to you?
42) In general, how would you rate your level of motivation?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
43) How would you rate your level of motivation to be physically active?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
44) How would you rate your ability to manage stress?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.

45) How would you rate your ability to manage your mood?	What does managing your mood mean to you?
46) How would you rate your ability to manage your weight?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
47) How would you rate your current life expectancy?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
48) How would you rate your ability to undertake everyday tasks that require some level of physical activity?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
49) How would you rate your overall quality of life?	What does the term quality of life mean to you?
50) How would you rate your current knowledge of the benefits of physical activity and exercise for health?	Can you repeat the question I just asked you in your own words? Probe for understanding of PA and exercise terms. Is there any difference?

Ending Probes:

How easy or difficult was it to complete the questionnaire?

"I don't think it was difficult at all"

Is there anything else you can tell me that you think will help with the development of the questionnaire?

No suggestions made.

THANK YOU

Appendix 9 Interviewer Probe Protocol 2

This questionnaire is being developed to help us understand aspects of quality of life that may be relevant or important to someone undertaking an exercise programme.

Please answer all the questions you are asked. If you are unsure about which response to give to a question, please choose the **ONE** that seems most appropriate. This can often be your first response. At some points during the questionnaire response options will change so we ask that you answer carefully.

We ask that you think about your life in the last two weeks when responding to the questions that will follow. Please **circle** one number only.

Before we move on to the actual questions, can you tell me what this introduction is

telling you?

1) How important to you is the affordability of leisure facilities in order to participate in a structured exercise programme?	What does the term structured exercise mean to you?
2) How much do you feel you have a choice regarding the exercise you undertake?	General probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
3) How much do you feel you have a choice regarding the time at which you exercise?	General probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
4) How much do you feel that lifestyle factors (e.g., transport, time, childcare, poor health & the weather) affect your ability to be physically active?	How easy or hard was it to choose an answer?
5) How much do you feel you compare yourself to others while exercising?	General probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
6) How much do you worry about exercising in an environment with others?	General probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.

7) How much do you worry about participating in structured exercise?	What does the term "structured exercise" mean to you?
8) How much do you worry about the prospect of exercising in a gym?	General probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
9) How much does any injury you may have prevent you from being physically active?	General probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking. Check the response scale is suitable for those who do not have any injury.
10) How much do you currently enjoy physical activity?	General probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
11) How much is social contact with others a part of your current lifestyle?	General probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
12) How important is it for you to manage your weight?	General probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
13) How much do you feel that you incorporate physical activity into your daily lifestyle?	What does the term "physical activity" mean to you? In what way does this differ to the word "exercise", if at all? "What sorts of things come to mind when you think of incorporating physical activity into your daily lifestyle?"
14) How well do you feel you adhere to eating habits that are beneficial to your health and any illness you may have?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.

15) How competitive are you?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking. Question the specificity of the question as appropriate (i.e., do respondents require more information regarding the context in which they are competitive)?
16) How determined are you?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking. Question the specificity of the question as appropriate (i.e., do respondents require more information regarding the context in which they are determined)?
17) How confident are you using gym equipment?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
18) How confident are you that you can exercise in a leisure centre with minimum support?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
19) How confident are you in your ability to participate in regular physical activity and exercise?	Can you repeat the question I just asked you in your own words? The question used the terms physical activity and exercise in the same question. Does that sound OK to you or would you choose something different?
20) How confident are you in your ability to undertake daily tasks that do not necessarily involve physical activity?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.

21) In general, how confident are you ar around other people?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
22) How confident are you in your own physical ability?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
23) How confident are you to join in activities with family and friends?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
24) How important is it to you right now to maintain a physically active lifestyle?	What does the term "physically active lifestyle" mean to you?
25) How confident are you, right now that you can maintain a physically active lifestyle?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
26) How motivated are you to maintain a physically active lifestyle?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
27) How likely are you to maintain a physically active lifestyle?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
28) In general, how much opportunity do you have to make time for yourself?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
29) How often do you experience physical pain?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.

30) How often do you feel you have to manage any symptoms of pain?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
31) How often do you feel you have to manage the symptoms of any illness you have, if any?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
32) How often do you feel you have achieved something in terms of your physical fitness?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
33) How would you rate your energy levels?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
34) How would you rate your outlook on life?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
35) How would you rate your level of self-esteem?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
36) How would you rate your physical health?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
37) How would you rate your mental health?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
38) How would you rate your sense of well-being?	What does the term “well-being” mean to you?
39) How would you rate the quality of support you receive from others to be physically active?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.

40) How would you rate the quality of your sleep?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
41) How would you rate your memory function?	What does the term “memory function” mean to you?
42) In general, how would you rate your level of motivation?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
43) How would you rate your level of motivation to be physically active?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
44) How would you rate your ability to manage stress?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
45) How would you rate your ability to manage your mood?	What does managing your mood mean to you?
46) How would you rate your ability to manage your weight?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
47) How would you rate your current life expectancy?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
48) How would you rate your ability to undertake everyday tasks that require some level of physical activity?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
49) How would you rate your overall quality of life?	What does the term quality of life mean to you?
50) How would you rate your current knowledge of the benefits of physical activity and exercise for health?	Can you repeat the question I just asked you in your own words? Probe for understanding of PA and exercise terms. Is there any difference?

Ending Probes:

How easy or difficult was it to complete the questionnaire?

“easy” “no problems for me”

Is there anything else you can tell me that you think will help with the development of the questionnaire?

No suggestions made.

THANK YOU

Appendix 10 Interviewer Probe Protocol 3

This questionnaire is being developed to help us understand aspects of quality of life that may be relevant or important to someone undertaking an exercise programme.

Please answer all the questions you are asked. If you are unsure about which response to give to a question, please choose the **ONE** that seems most appropriate. This can often be your first response. At some points during the questionnaire response options will change so we ask that you answer carefully.

We ask that you think about your life in the last two weeks when responding to the questions that will follow. Please **circle** one number only.

Before we move on to the actual questions, can you tell me what this introduction is telling you?

1) How important to you is the affordability of leisure facilities in order to participate in a structured exercise programme?	What does the term structured exercise mean to you?
2) How much do you feel you have a choice regarding the exercise you undertake?	General probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
3) How much do you feel you have a choice regarding the time at which you exercise?	General probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
4) How much do you feel that lifestyle factors (e.g., transport, time, childcare, poor health & the weather) affect your ability to be physically active?	How easy or hard was it to choose an answer? You will notice that I provided you with some examples of lifestyle factors (transport, time, childcare, poor health & the weather) did these examples help you or make it more difficult to answer the question?

5) How much do you feel you compare yourself to others while exercising?	General probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
6) How much do you worry about exercising in an environment with others?	General probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
7) How much do you worry about participating in structured exercise?	What does the term "structured exercise" mean to you? General probes if gathered enough info from probing understanding of the term in Q1
8) How much do you worry about the prospect of exercising in a gym?	General probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
9) How much does any injury you may have prevent you from being physically active?	General probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking. Check the response scale is suitable for those who do not have any injury.
10) How much do you currently enjoy physical activity?	General probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
11) How much is social contact with others a part of your current lifestyle?	General probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
12) How important is it for you to manage your weight?	General probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.

13) How much do you feel that you incorporate physical activity into your daily lifestyle?	What does the term "physical activity" mean to you? In what way does this differ to the word "exercise", if at all? "What sorts of things come to mind when you think of incorporating physical activity into your daily lifestyle?"
14) How well do you feel you adhere to eating habits that are beneficial to your health and any illness you may have?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
15) How competitive are you?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking. Question the specificity of the question as appropriate (i.e., do respondents require more information regarding the context in which they are competitive)?
16) How determined are you?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking. Question the specificity of the question as appropriate (i.e., do respondents require more information regarding the context in which they are determined)?
17) How confident are you using gym equipment?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
18) How confident are you that you can exercise in a leisure centre with minimum support?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.

19) How confident are you in your ability to participate in regular physical activity and exercise?	Can you repeat the question I just asked you in your own words? The question used the terms physical activity and exercise in the same question. Does that sound OK to you or would you choose something different? If the terms are described as having different meanings, probe for how they came to an answer (i.e., did they focus on one more than the other)?
20) How confident are you in your ability to undertake daily tasks that do not necessarily involve physical activity?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
21) In general, how confident are you ar around other people?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
22) How confident are you in your own physical ability?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
23) How confident are you to join in activities with family and friends?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
24) How important is it to you right now to maintain a physically active lifestyle?	What does the term “physically active lifestyle” mean to you?
25) How confident are you, right now that you can maintain a physically active lifestyle?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
26) How motivated are you to maintain a physically active lifestyle?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
27) How likely are you to maintain a physically active lifestyle?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.

28) In general, how much opportunity do you have to make time for yourself?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
29) How often do you experience physical pain?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
30) How often do you feel you have to manage any symptoms of pain?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
31) How often do you feel you have to manage the symptoms of any illness you have, if any?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
32) How often do you feel you have achieved something in terms of your physical fitness?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
33) How would you rate your energy levels?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
34) How would you rate your outlook on life?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
35) How would you rate your level of self-esteem?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
36) How would you rate your physical health?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
37) How would you rate your mental health?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.

38) How would you rate your sense of well-being?	What does the term "well-being" mean to you?
39) How would you rate the quality of support you receive from others to be physically active?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
40) How would you rate the quality of your sleep?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
41) How would you rate your memory function?	What does the term "memory function" mean to you?
42) In general, how would you rate your level of motivation?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
43) How would you rate your level of motivation to be physically active?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
44) How would you rate your ability to manage stress?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
45) How would you rate your ability to manage your mood?	What does managing your mood mean to you?
46) How would you rate your ability to manage your weight?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
47) How would you rate your current life expectancy?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.
48) How would you rate your ability to undertake everyday tasks that require some level of physical activity?	General Probes: How did you arrive at that answer? Was that easy or hard to answer? I noticed that you hesitated. Tell me what you were thinking.

49) How would you rate your overall quality of life?	What does the term quality of life mean to you? What kind of things come to mind when you hear the term quality of life?
50) How would you rate your current knowledge of the benefits of physical activity and exercise for health?	Can you repeat the question I just asked you in your own words? Probe for understanding of PA and exercise terms. Is there any difference?

Ending Probes:

How easy or difficult was it to complete the questionnaire?

"yes very easy no difficulties, it makes sense to me"

Is there anything else you can tell me that you think will help with the development of the questionnaire?

No suggestions made.

THANK YOU

Appendix 11 Pre-Testing Outcome Report: Self Complete

This questionnaire is being developed to help us understand aspects of quality of life that may be relevant or important to someone undertaking an exercise programme.

Please answer all the questions you are asked. If you are unsure about which response to give to a question, please choose the **ONE** that seems most appropriate. This can often be your first response. At some points during the questionnaire response options will change so we ask that you answer carefully.

We ask that you think about your life in the last two weeks when responding to the questions that will follow. Please **circle** one number only.

Before we move on to the actual questions, can you tell me what this introduction is telling you?

No difficulties understanding the instructions were reported.

Interpretations included:

"Read the questions and actually circle which one thing is the appropriate answer to it, but to actually think about the last two weeks of my lifestyle to see what corresponds to that"

"Choose carefully, one answer per question"

"Read each question and circle an answer"

1) How important to you is the affordability of leisure facilities in order to participate in a structured exercise programme?	No problems understanding the question or response reported. Structured exercise was understood as exercise with support/exercise with a leader.

2) How much do you feel you have a choice regarding the exercise you undertake?	No problems understanding the question or response reported. Choice of exercise was more commonly understood as the individual exercises done as part of a circuit class for those attending this class as opposed to the generic mode per sae. Consider amending question to 'type of exercise'.
3) How much do you feel you have a choice regarding the time at which you exercise?	No problems understanding the question or response reported.
4) How much do you feel that lifestyle factors (e.g., transport, time, childcare, poor health & the weather) affect your ability to be physically active?	No problems understanding the question or response reported. Respondents were able to verbalise the lifestyle factor that affected their ability to be physically active, if any. Probes provided one suggestion to have a question for each example and one suggestion to remove the examples. However, all other respondents reported that the examples were helpful in assisting a response. The term 'childcare' was also reported as suitable for the care of a grandchild in addition to parental responsibilities.
5) How much do you feel you compare yourself to others while exercising?	No problems understanding the question or response reported. One suggestion to remove "you feel".
6) How much do you worry about exercising in an environment with others?	No problems understanding the question or response reported.
7) How much do you worry about participating in structured exercise?	No problems understanding the question or response reported.
8) How much do you worry about the prospect of exercising in a gym?	No problems understanding the question or response reported. As per qualitative data, support was reported as an important mediator for the level of worry experienced.

9) How much does any injury you may have prevent you from being physically active?	No problems understanding the question or response reported. The term 'injury' may also be sufficient in referring to any post-operative discomfort. Suggestion to include an apostrophe following 'may have'.
10) How much do you currently enjoy physical activity?	No problems understanding the question or response reported.
11) How much is social contact with others a part of your current lifestyle?	No problems understanding the question or response reported. One suggestion to use a comma after "others".
12) How important is it for you to manage your weight?	No problems understanding the question or response reported. The response 'not at all' is appropriate for those who have no difficulties managing their weight.
13) How much do you feel that you incorporate physical activity into your daily lifestyle?	No problems understanding the question or response reported.
14) How well do you feel you adhere to eating habits that are beneficial to your health and any illness you may have?	One suggestion to substitute the word "adhere" for "stick-to" or something similar. Consider substituting 'adhere' for 'maintain'.
15) How competitive are you?	No problems understanding the question or response reported.
16) How determined are you?	No problems understanding the question or response reported.
17) How confident are you using gym equipment?	No problems understanding the question or response reported. For those who had not used gym equipment "not at all" was an appropriate response and posed no difficulty.
18) How confident are you that you can exercise in a leisure centre with minimum support?	One participant reported that clarification may be required regarding the source of support (i.e., family & friends or instructor).

19) How confident are you in your ability to participate in regular physical activity and exercise?	No problems understanding the question or response reported. "physical activity" and "exercise" in the same question posed no difficulties.
20) How confident are you in your ability to undertake daily tasks that do not necessarily involve physical activity?	No problems understanding the question or response reported.
21) In general, how confident are you around around other people?	No problems understanding the question or response reported.
22) How confident are you in your own physical ability?	No problems understanding the question or response reported.
23) How confident are you to join in activities with family and friends?	No problems understanding the question or response reported.
24) How important is it to you right now to maintain a physically active lifestyle?	No problems understanding the question or response reported.
25) How confident are you, right now that you can maintain a physically active lifestyle?	No problems understanding the question or response reported. One suggestion to insert a comma after "right now".
26) How motivated are you to maintain a physically active lifestyle?	No problems understanding the question or response reported.
27) How likely are you to maintain a physically active lifestyle?	No problems understanding the question or response reported. Response 3 (moderately) may be perceived as a neutral response as one participant stated "I'll sit on the fence".
28) In general, how much opportunity do you have to make time for yourself?	No problems understanding the question or response reported.
29) How often do you experience physical pain?	No problems understanding the question or response reported.

30) How often do you feel you have to manage any symptoms of pain?	No problems understanding the question or response reported. 'Management' may include the use of painkillers.
31) How often do you feel you have to manage the symptoms of any illness you have, if any?	No problems understanding the question or response reported. Three participants suggested removing "if any".
32) How often do you feel you have achieved something in terms of your physical fitness?	No problems understanding the question or response reported.
33) How would you rate your energy levels?	No problems understanding the question or response reported.
34) How would you rate your outlook on life?	No problems understanding the question or response reported.
35) How would you rate your level of self-esteem?	No problems understanding the question or response reported.
36) How would you rate your physical health?	No problems understanding the question or response reported.
37) How would you rate your mental health?	No problems understanding the question or response reported.
38) How would you rate your sense of well-being?	No problems understanding the question or response reported.
39) How would you rate the quality of support you receive from others to be physically active?	No problems understanding the question or response reported. "Others" perceived as family, work colleagues and leisure centre staff.
40) How would you rate the quality of your sleep?	No problems understanding the question or response reported.
41) How would you rate your memory function?	No problems understanding the question or response reported.
42) In general, how would you rate your level of motivation?	No problems understanding the question or response reported.
43) How would you rate your level of motivation to be physically active?	No problems understanding the question or response reported.

44) How would you rate your ability to manage stress?	No problems understanding the question or response reported.
45) How would you rate your ability to manage your mood?	No problems understanding the question or response reported.
46) How would you rate your ability to manage your weight?	No problems understanding the question or response reported. The response 'very good' is sufficient to adequately reflect the responses of those who have no difficulty with weight management.
47) How would you rate your current life expectancy?	No problems understanding the question or response reported. One participant reported considering other risk factors other than physical inactivity before responding "smoking and drinking and food".
48) How would you rate your ability to undertake everyday tasks that require some level of physical activity?	No problems understanding the question or response reported. Examples of activities reported include gardening and walking.
49) How would you rate your overall quality of life?	No problems understanding the question or response reported.
50) How would you rate your current knowledge of the benefits of physical activity and exercise for health?	No problems understanding the question or response reported. The terms "physical activity" and "exercise" in the same question posed no reported difficulties. The terms are understood

Ending Probes:

How easy or difficult was it to complete the questionnaire?

"Quite easy, quite happy"

"Overall I didn't have any problem doing it"

"Easy"

Is there anything else you can tell me that you think will help with the development of the questionnaire?

No additional suggestions made.

Appendix 12 Pre-Testing Outcome Report: Interview-Administered

This questionnaire is being developed to help us understand aspects of quality of life that may be relevant or important to someone undertaking an exercise programme.

Please answer all the questions you are asked. If you are unsure about which response to give to a question, please choose the **ONE** that seems most appropriate. This can often be your first response. At some points during the questionnaire response options will change so we ask that you answer carefully.

We ask that you think about your life in the last two weeks when responding to the questions that will follow. Please **circle** one number only.

Before we move on to the actual questions, can you tell me what this introduction is

telling you?

No difficulties understanding the instructions were reported.

1) How important to you is the affordability of leisure facilities in order to participate in a structured exercise programme?	No problems understanding the question or response reported. Structured exercise was understood as exercise that was 'supervised' or 'organised'.
2) How much do you feel you have a choice regarding the exercise you undertake?	No problems understanding the question or response reported. Choice of exercise during classes vs. type needs clarification. Consider adding 'type of exercise' to the question.
3) How much do you feel you have a choice regarding the time at which you exercise?	No problems understanding the question or response reported. Limited choice reported by those attending classes.
4) How much do you feel that lifestyle factors (e.g., transport, time, childcare, poor health & the weather) affect your ability to be physically active?	No problems understanding the question or response reported. Respondents were able to verbalise the lifestyle factor that affected their ability to be physically active, if any. Examples were helpful in assisting a response. The term 'childcare' was also reported as suitable for the care of a grandchild in addition to parental responsibilities.

5) How much do you feel you compare yourself to others while Exercising?	No problems understanding the question or response reported. One suggestion to remove "you feel".
6) How much do you worry about exercising in an environment with others?	No problems understanding the question or response reported.
7) How much do you worry about participating in structured exercise?	No problems understanding the question or response reported.
8) How much do you worry about the prospect of exercising in a gym?	No problems understanding the question or response reported. Sometimes 'gym' is referred to as a 'fitness suite'.
9) How much does any injury you may have prevent you from being physically active?	No problems understanding the question or response reported.
10) How much do you currently enjoy physical activity?	No problems understanding the question or response reported. Examples of physical activity included walking and gardening.
11) How much is social contact with others a part of your current lifestyle?	No problems understanding the question or response reported.
12) How important is it for you to manage your weight?	No problems understanding the question or response reported.
13) How much do you feel that you incorporate physical activity into your daily lifestyle?	No problems understanding the question or response reported.
14) How well do you feel you adhere to eating habits that are beneficial to your health and any illness you may have?	No problems understanding the question or response reported.
15) How competitive are you?	Two participant's reported that clarification may be required as to the context of this question (i.e., generally or with regard to exercise).
16) How determined are you?	Two participant's reported that clarification may be required as to the context of this question (i.e., generally or with regard to exercise).

17) How confident are you using gym equipment?	No problems understanding the question or response reported.
18) How confident are you that you can exercise in a leisure centre with minimum support?	No problems understanding the question or response reported. One participant suggested adding 'if you had to'.
19) How confident are you in your ability to participate in regular physical activity and exercise?	No problems understanding the question or response reported. "physical activity" and "exercise" in the same question posed no difficulties.
20) How confident are you in your ability to undertake daily tasks that do not necessarily involve physical activity?	No problems understanding the question or response reported.
21) In general, how confident are you ar around other people?	No problems understanding the question or response reported.
22) How confident are you in your own physical ability?	No problems understanding the question or response reported. Physical ability interpreted as "the ability to be able to do things". The question also helped respondents consider their limitations as per qualitative data.
23) How confident are you to join in activities with family and friends?	No problems understanding the question or response reported.
24) How important is it to you right now to maintain a physically active lifestyle?	No problems understanding the question or response reported.
25) How confident are you, right now that you can maintain a physically active lifestyle?	No problems understanding the question or response reported.
26) How motivated are you to maintain a physically active lifestyle?	No problems understanding the question or response reported.

27) How likely are you to maintain a physically active lifestyle?	No problems understanding the question or response reported.

28) In general, how much opportunity do you have to make time for yourself?	No problems understanding the question or response reported.
29) How often do you experience physical pain?	No problems understanding the question or response reported.
30) How often do you feel you have to manage any symptoms of pain?	No problems understanding the question or response.
31) How often do you feel you have to manage the symptoms of any illness you have, if any?	No problems understanding the question or response. Questioning probe suggested removing 'if any'.
32) How often do you feel you have achieved something in terms of your physical fitness?	No problems understanding the question or response reported.
33) How would you rate your energy levels?	No problems understanding the question or response reported.
34) How would you rate your outlook on life?	No problems understanding the question or response reported.
35) How would you rate your level of self-esteem?	No problems understanding the question or response reported.
36) How would you rate your physical health?	No problems understanding the question or response reported.
37) How would you rate your mental health?	No problems understanding the question or response reported.
38) How would you rate your sense of well-being?	No problems understanding the question or response reported. One participant reported interpreting well-being as 'contentment'.
39) How would you rate the quality of support you receive from others to be physically active?	No problems understanding the question or response.
40) How would you rate the quality of your sleep?	No problems understanding the question or response reported.

41) How would you rate your memory function?	No problems understanding the question or response reported. Respondents considered both their short and long-term memory before responding.
42) In general, how would you rate your level of motivation?	No problems understanding the question or response reported. Two suggestions to make it more clear as to whether this is in general or with specific reference to physical activity.
43) How would you rate your level of motivation to be physically active?	No problems understanding the question or response reported.
44) How would you rate your ability to manage stress?	No problems understanding the question or response reported.
45) How would you rate your ability to manage your mood?	No problems understanding the question or response reported.
46) How would you rate your ability to manage your weight?	No problems understanding the question or response reported. No problems responding when weight management is not a primary health issue.
47) How would you rate your current life expectancy?	No problems understanding the question or response reported.
48) How would you rate your ability to undertake everyday tasks that require some level of physical activity?	No problems understanding the question or response reported
49) How would you rate your overall quality of life?	No problems understanding the question or response reported.
50) How would you rate your current knowledge of the benefits of physical activity and exercise for health?	No problems understanding the question or response reported. The terms "physical activity" and "exercise" in the same question posed no reported difficulties. The terms are understood. Two participants reported the role of the media in contributing to conflicting health messages (e.g., alcohol and salt consumption).

Ending Probes:

How easy or difficult was it to complete the questionnaire?

“very easy” “no problems” “it’s very good”

Is there anything else you can tell me that you think will help with the development of the questionnaire? No additional suggestions made.

Appendix 13 Pre-Testing Outcome Report: Telephone Administered

This questionnaire is being developed to help us understand aspects of quality of life that may be relevant or important to someone undertaking an exercise programme.

Please answer all the questions you are asked. If you are unsure about which response to give to a question, please choose the **ONE** that seems most appropriate. This can often be your first response. At some points during the questionnaire response options will change so we ask that you answer carefully.

We ask that you think about your life in the last two weeks when responding to the questions that will follow. Please **circle** one number only.

Before we move on to the actual questions, can you tell me what this introduction is telling you?

No difficulties understanding the instructions were reported.

Responses to the instructions included that they were "clear", "very clear" and there was "no difficulty at all" in understanding what was being requested.

1) How important to you is the affordability of leisure facilities in order to participate in a structured exercise programme?	No problems understanding the question or response reported. Structured exercise was understood as 'activities undertaken at a particular time, and exercise that was 'supervised' or 'organised'.
2) How much do you feel you have a choice regarding the exercise you undertake?	No problems understanding the question or response reported. For those participating in classes, exercise undertaken could refer to the different exercises delivered in the class.
3) How much do you feel you have a choice regarding the time at which you exercise?	No problems understanding the question or response reported.

4) How much do you feel that lifestyle factors (e.g., transport, time, childcare, poor health & the weather) affect your ability to be physically active?	No problems understanding the question or response reported. Respondents were able to verbalise the lifestyle factor that affected their ability to be physically active, if any. Examples were helpful in assisting a response.
5) How much do you feel you compare yourself to others while exercising?	No problems understanding the question or response reported.
6) How much do you worry about exercising in an environment with others?	No problems understanding the question or response reported.
7) How much do you worry about participating in structured exercise?	No problems understanding the question or response reported.
8) How much do you worry about the prospect of exercising in a gym?	No problems understanding the question or response reported.
9) How much does any injury you may have prevent you from being physically active?	No problems understanding the question or response reported. The response option 'not at all' is suitable for those reporting no injury.
10) How much do you currently enjoy physical activity?	No problems understanding the question or response reported. Examples of physical activity included walking and gardening.
11) How much is social contact with others a part of your current lifestyle?	No problems understanding the question or response reported.
12) How important is it for you to manage your weight?	No problems understanding the question or response reported.
13) How much do you feel that you incorporate physical activity into your daily lifestyle?	No problems understanding the question or response reported. One participant reported including exercises at home into her response choice.

14) How well do you feel you adhere to eating habits that are beneficial to your health and any illness you may have?	No problems understanding the question or response reported.
15) How competitive are you?	Two participant's reported that clarification may be required as to the context of this question (i.e., generally or with regard to exercise).
16) How determined are you?	Two participant's reported that clarification may be required as to the context of this question (i.e., generally or with regard to exercise).
17) How confident are you using gym equipment?	No problems understanding the question or response reported. One participant reported that support from an instructor was important as per qualitative data findings.
18) How confident are you that you can exercise in a leisure centre with minimum support?	No problems understanding the question or response reported.
19) How confident are you in your ability to participate in regular physical activity and exercise?	No problems understanding the question or response reported. "physical activity" and "exercise" in the same question posed no difficulties.
20) How confident are you in your ability to undertake daily tasks that do not necessarily involve physical activity?	No problems understanding the question or response reported.
21) In general, how confident are you around around other people?	No problems understanding the question or response reported.
22) How confident are you ir own physical ability?	No problems understanding the question or response reported. Physical ability interpreted as "physically able to do things".
23) How confident are you to join in Activities with family and friends?	No problems understanding the question or response reported. One participant reported not having any family. However, this posed no difficulties in responding.

24) How important is it to you right now to maintain a physically active lifestyle?	No problems understanding the question or response reported. Physically active lifestyle was often referred to as “movement” and “active daily living”
25) How confident are you, right now that you can maintain a physically active lifestyle?	No problems understanding the question or response reported.
26) How motivated are you to maintain a physically lifestyle?	No problems understanding the question or response reported.
27) How likely are you to maintain a physically active lifestyle?	No problems understanding the question or response reported.

28) In general, how much opportunity do you have to make time for yourself?	No problems understanding the question or response reported.
29) How often do you experience physical pain?	No problems understanding the question or response reported.
30) How often do you feel you have to manage any symptoms of pain?	No problems understanding the question or response.
31) How often do you feel you have to manage the symptoms of any illness you have, if any?	No problems understanding the question or response. Alcoholism was included in the response of one participant.
32) How often do you feel you have achieved something in terms of your physical fitness?	No problems understanding the question or response reported.
33) How would you rate your energy levels?	No problems understanding the question or response reported.
34) How would you rate your outlook on life?	No problems understanding the question or response reported.

35) How would you rate your level of self-esteem?	No problems understanding the question or response reported. One participant reported that there were a number of factors to consider before responding. However, reported that the question was worded adequately and there were no difficulties with interpretation.
36) How would you rate your physical health?	No problems understanding the question or response reported.
37) How would you rate your mental health?	No problems understanding the question or response reported.
38) How would you rate your sense of well-being?	No problems understanding the question or response reported. Well-being was interpreted as "being in control of a situation and feeling good" and "feeling good about yourself"
39) How would you rate the quality of support you receive from others to be physically active?	No problems understanding the question or response.
40) How would you rate the quality of your sleep?	No problems understanding the question or response reported.
41) How would you rate your memory function?	No problems understanding the question or response reported.
42) In general, how would you rate your level of motivation?	No problems understanding the question or response reported. One suggestions to make it more clear as to whether this is in general or with specific reference to physical activity.
43) How would you rate your level of motivation to be physically active?	No problems understanding the question or response reported.
44) How would you rate your ability to manage stress?	No problems understanding the question or response reported.
45) How would you rate your ability to manage your mood?	No problems understanding the question or response reported.
46) How would you rate your ability to manage your weight?	No problems understanding the question or response reported. No problems responding when weight management is not a primary health issue.

47) How would you rate your current life expectancy?	No problems understanding the question or response reported.
48) How would you rate your ability to undertake everyday tasks that require some level of physical activity?	No problems understanding the question or response reported
49) How would you rate your overall quality of life?	No problems understanding the question or response reported. Quality of life was described as related to the amount of effort that was put into things, “making the most of life”, “appreciating things”. One participant stated “how you live each day and if you are satisfied with the quality of each day that you have”.
50) How would you rate your current knowledge of the benefits of physical activity and exercise for health?	No problems understanding the question or response reported. The terms “physical activity” and “exercise” in the same question posed no reported difficulties. The terms are understood. One participant reported that physical activity may be more general, while exercise is more structured. One participant reported that “physical activity can be exercise”.

Ending Probes:

How easy or difficult was it to complete the questionnaire?

Participants reported “no difficulty at all” with completing the questionnaire and that it was “easy” to understand and complete. One participant stated that the questions are “particularly relevant to older people”

Is there anything else you can tell me that you think will help with the development of the questionnaire?

No additional suggestions made.

Appendix 14 Test Questionnaire

Participant Name_____ Code_office use____

Date of birth_____

Gender (please tick) Male ☐ Female ☐

Primary Reason(s) for Referral_____
_____

Date Questionnaire was Administered_____/_____/_____

Stage of Referral when Questionnaire was Administered (please tick):

Exercise Programme Start ☐
Mid Exercise Programme ☐
Post Exercise Programme ☐

Scheme Name/Location_____

Ethnicity (please tick): White ☐ Mixed ☐
Indian ☐ Pakistani ☐ Bangladeshi ☐
Other Asian ☐ Black Caribbean ☐
Black African ☐ Other Black ☐ Chinese ☐
Other Ethnic Group ☐

This questionnaire is being developed to help us understand aspects of quality of life that may be relevant or important to someone undertaking an exercise programme.

The first set of questions you will be asked are the ones under development for the exercise quality of life questionnaire. You will also be asked questions from an existing mood questionnaire and also an existing general quality of life questionnaire that will help us to test the performance of the questionnaire under development. You will be informed when each questionnaire begins.

Please answer all the questions you are asked. If you are unsure about which response to give to a question, **please choose the ONE** that seems most appropriate. This can often be your first response. At some points during the questionnaire response options will change so we ask that you answer carefully.

The following questions relate to the exercise quality of life questionnaire under development. We ask that you think about your **life in the last two weeks** when responding to the questions that will follow. Please **circle** one number only.

	1 Not at all	**2 Not much**	**3 Moderately**	**4 A great deal**	**5 Completely**
1) How important to you is the affordability of leisure facilities in order to participate in a structured exercise programme?	1	2	3	4	5
2) To what extent do you feel you have a choice regarding the type of exercise you undertake?	1	2	3	4	5
3) To what extent do you feel you have a choice regarding the time at which you exercise?	1	2	3	4	5
4) How much do you feel that lifestyle factors (e.g., transport, time, childcare, poor health & the weather) affect your ability to be physically active?	1	2	3	4	5
5) How much do you compare yourself to others while exercising?	1	2	3	4	5

	1 Not at all	**2 Not much**	**3 Moderately**	**4 A great deal**	**5 Completely**
6) To what extent do you worry about exercising in an environment with others?	1	2	3	4	5
7) To what extent do you worry about participating in structured exercise?	1	2	3	4	5
8) To what extent do you worry about the prospect of exercising in a fitness suite/gym?	1	2	3	4	5
9) How much does any injury you may have prevent you from being physically active?	1	2	3	4	5
10) How much do you currently enjoy physical activity?	1	2	3	4	5
11) How much is social contact with others a part of your current lifestyle?	1	2	3	4	5
12) How important is it for you to manage your weight?	1	2	3	4	5
13) How much do you feel that you incorporate physical activity into your daily lifestyle?	1	2	3	4	5

14) How well do you feel you maintain eating habits that are beneficial to your health and any illness you may have?	1	2	3	4	5
15) In terms of exercise how competitive are you?	1	2	3	4	5
16) In terms of exercise how determined are you?	1	2	3	4	5
17) How confident are you using gym equipment?	1	2	3	4	5
18) How confident are you that you can exercise in a leisure centre with minimum support?	1	2	3	4	5
19) How confident are you in your ability to participate in regular physical activity and exercise?	1	2	3	4	5
20) How confident are you in your ability to undertake daily tasks that do not necessarily involve physical activity?	1	2	3	4	5
21) In general, confident are you around around other people?	1	2	3	4	5

22) How confident are you in your own physical ability?	1	2	3	4	5
23) How confident are you to join in activities with family and friends?	1	2	3	4	5
24) How important is it you right now to maintain a physically active lifestyle?	1	2	3	4	5
25) How confident are you, right now that you can maintain a physically active lifestyle?	1	2	3	4	5
26) How motivated are you to maintain a physically active lifestyle?	1	2	3	4	5
27) How likely are you to maintain a physically active lifestyle?	1	2	3	4	5
	1 Never	**2 Seldom**	**3 Quite often**	**4 Very often**	**5 Always**
28) In general, how much opportunity do you have to make time for yourself?	1	2	3	4	5
29) How often do you experience physical pain?	1	2	3	4	5

30) How often do you feel you have to manage any symptoms of pain?	1	2	3	4	5
31) How often do you feel you have to manage the symptoms of any illness you have?	1	2	3	4	5
32) How often do you feel you have achieved something in terms of your physical fitness?	1	2	3	4	5
	1 Very Poor	**2 Poor**	**3 Neither good nor poor**	**4 Good**	**5 Very good**
33) How would you rate your energy levels?	1	2	3	4	5
34) How would you rate your outlook on life?	1	2	3	4	5
35) How would you rate your level of self-esteem?	1	2	3	4	5
36) How would you rate your physical health?	1	2	3	4	5
37) How would you rate your mental health?	1	2	3	4	5

38) How would you rate your sense of well-being?	1	2	3	4	5
39) How would you rate the quality of support you receive from leisure staff to be physically active?	1	2	3	4	5
40) How would you rate the quality of support you receive from family and friends to be physically active?	1	2	3	4	5
41) How would you rate the quality of your sleep?	1	2	3	4	5
42) How would you rate your memory function?	1	2	3	4	5
43) How would you rate your level of motivation to be physically active?	1	2	3	4	5
44) How would you rate your ability to manage stress?	1	2	3	4	5
45) How would you rate your ability to manage your mood?	1	2	3	4	5
46) How would you rate your ability to manage your weight?	1	2	3	4	5

47) How would you rate your current life expectancy?	1	2	3	4	5
48) How would you rate your ability to undertake everyday tasks that require some level of physical activity?	1	2	3	4	5
49) How would you rate your overall quality of life?	1	2	3	4	5
50) How would you rate your current knowledge of the benefits of physical activity and exercise for health?	1	2	3	4	5

Thank you for taking the time to complete this questionnaire

Appendix 15 NRES Participant Information Sheet: Questionnaire Validation

Part 1

You are being invited to take part in a research study. Before you decide it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully. Talk to others about the study if you wish.

The number of exercise referral programmes just like the scheme you have been referred into are growing. To date, there is no standardised method of evaluating these programmes which is essential for their continuation. It is also important that there is adequate opportunity for the participants to explore their experiences in order that the reasons for continuing to be physically active for health may be understood and individuals have an appropriate opportunity to explore their personal benefits and barriers to a more active lifestyle

This research study aims to develop a quality of life questionnaire that is specifically designed to measure those items of importance to people participating in an exercise referral scheme. You are being requested to contribute to the development of the questionnaire.

Please ask your instructor or the primary investigator if you would like more information. Take time to decide whether or not you wish to take part.

What is the purpose of the study?

This research is primarily for educational purposes. The primary researcher is undertaking a PhD at Nottingham Trent University. It is expected that the outcome of this research will also help other referral schemes to assess the needs of their clients and evaluate their programmes more effectively.

Why have I been chosen?

You have been chosen to complete the questionnaire because you have been referred for exercise and you are in an optimum position to convey your thoughts and feelings regarding how well the questionnaire under development captures what is important to people like yourself.

Do I have to take part?

No. It is up to you to decide whether or not to take part. If you do, you will be given this information sheet to keep and be asked to sign a consent form. You are still free to withdraw at any time and without giving a reason. A decision to withdraw at any time, or a decision not to take part, will not affect the standard of care you receive. In addition, if you decide that you do not with the questionnaire you complete to be included in the study you have until April 30th to notify the research team.

What will happen to me if I take part?

- You will be asked to complete a questionnaire for exercise referral schemes that is under development.
- You will also be asked to complete two additional questionnaires that have been added to the one under development. These two additional questionnaires have been added to the test measure to make a single questionnaire.
- If the results of the research are shared at academic conferences, journals and or public media, you will not be identified by name or otherwise unless consent is given by you specifically following the completion of the research.
- There is no long-term monitoring or follow up planned specifically. However, if you wish to contact the research team during the study period (planned to end December 2009), you are welcome to do so.

What do I have to do?

If you decide to complete the questionnaire, after completing a consent form you will be asked to complete the questionnaire under development and two additional questionnaires in a single sitting. For some participants, we may ask that you complete the questionnaire again up to seven days later. You will be informed by someone at the exercise scheme that you attend if this is the case. Your completed questionnaire(s) will then be mailed to the primary researcher at Nottingham Trent University.

What are the possible benefits of taking part?

There is no direct clinical benefit from your participation. However, your participation is likely to assist the future development and evaluation procedures of exercise referral schemes and the experiences of those referred into them.

What happens when the research study stops?

When the research stops, the results of the research may be published in academic journals, local authority and Primary Care Trust publications and communicated at appropriate conferences. If you wish to know more about the course of the research, you may contact the research team (the contact details of whom are given below).

What if there is a problem?

Any complaint about the way you have been dealt with during the study or any possible harm you might suffer will be addressed. The detailed information regarding this is given in Part Two where you will also find a contact number for complaints.

Will my taking part in the study be kept confidential?

Yes. All the information regarding your participation in this study will be kept confidential. The details are included in part two.

Contact details

The primary investigator of this research is Charlotte Hilton. You may contact her or the primary research supervisor during the hours of 9.30am and 5.30pm.

Address:
Charlotte Hilton
Room 204
Erasmus Darwin
Sport Science
School of Science and Technology
Nottingham Trent University
Nottingham
NG11 8NS
Telephone: 07939247921
Email: charlotte.hilton@ntu.ac.uk

Part Two

What will happen if I don't want to carry on with the study?

If you withdraw from the study, we will destroy all your identifiable data, however, we will need to use the data collected up until your withdrawal. The final date that you can withdraw from the study is April 30th 2009.

What if there is a problem?

Should you experience a problem either during or following your participation in the focus group, you may contact the primary investigator. Alternatively, you may wish to contact a member of the research team:

Primary Research Supervisor:
Dr. Antoinette Minniti
School of Science and Technology
Nottingham Trent University
Nottingham
NG11 8NS
Telephone: 0115 8483918

Complaints

If you have a concern about any aspect of this study, you should ask to speak with the researchers (contact details above) who will do their best to answer your questions. If you remain unhappy and wish to complain formally, you can do this through the Nottingham Trent University complaints procedure. Details can be obtained from the university.

Harm

In the event that something goes wrong and you are harmed during the research study there are no special compensation arrangements. If you are harmed and this is due to someone's negligence then you may have grounds for a legal action for compensation against Nottingham Trent University but you may have to pay your legal costs. The normal National Health Service complaints mechanisms will still be available to you.

Will my taking part in this study be kept confidential?

Yes. The data for this study will be collected by your responses to the questions contained within the questionnaire. Subsequently, these responses will be entered into a statistical software package and at this point your responses will be coded and remain completely anonymous. The primary researcher and Dr Minniti are the only individuals who will have access to this data before it is coded (made anonymous) and also thereafter.

The custodian for the data is the primary investigator and their research supervisor (Dr. Antoinette Minniti). The data will be stored in locked filing cabinets to which only the primary investigator and their research supervisor will have access. The data will not be retained for use in future studies. The data will be retained for up to 3 years and disposed of securely.

What will happen to the results of the research study?

It is intended to communicate the results of the study overall within academic journals and conferences. There may also be an opportunity to communicate the findings within Primary Care Trust (PCT) publications. You will not be identified in any presentation or publication unless you have consented to release such information.

Who is organising and funding the research?

Nottingham Trent University is part-funding the research by paying the academic fees of the primary researcher. There is a small stationery budget and all other costs are met by the primary researcher.

Who has reviewed the study?

This study was given a favourable ethical opinion for conduct in the NHS by an NHS Research Ethics Committee (REC) and Nottingham Trent University ethics.

Thank you for considering to take part and taking the time to read this information.

NOTTINGHAM TRENT UNIVERSITY

Appendix 16 Questionnaire Validation Information for Participants

You have been asked to participate in this research because you have been referred for exercise to improve your health.

The complete research study aims to develop a questionnaire to help us understand aspects of quality of life that may be relevant or important to someone undertaking an exercise programme.

If you decide to take part, you will be asked to complete a questionnaire that is under development. You will also be asked to complete two additional questionnaires that have been added to the one under development. These two additional questionnaires have been added to the test measure to make a single questionnaire.

Your completed questionnaire will be returned to the primary researcher and your responses will help to develop a final measure that will be reliable and valid for use within exercise referral settings.

If you decide that you do not wish your completed questionnaire to be included in the study, you have until August 7th 2009 to inform the research team (details below).

If you have any questions once the session is completed, you may contact:

Charlotte Hilton:
Email: charlotte.hilton@ntu.ac.uk
Telephone: 07939247921

Toni Minniti:
Email: antoinette.minniti@ntu.ac.uk
Telephone: 0115 88483918

NOTTINGHAM TRENT UNIVERSITY

Appendix 17 Cover Letter to Recruitment Sites

Validation of an Exercise Referral Sensitive Quality of Life Measure

THANK YOU for agreeing to support this research that aims to validate a quality of life (QoL) measure specifically for use within exercise referral settings.

Please request that participants complete the questionnaire on two occasions, seven days apart in the first instance. However, if participants are only willing to complete on one occasion, we would like them to do so.

You should have all the resources you need including the questionnaire, consent form and information for participants (short and long version). Please print and use copies as you need them.

Important note:
Please consider all scheme participants from week one through to those that have been exercising for up to 12 months since their referral. Three weeks will be allocated to schemes in order to complete as many questionnaires as is feasible for the size of your scheme and resources. Please return the questionnaires weekly as they are completed. If you are administering the questionnaire twice (7 days apart) please send both questionnaires for each person together.

Once the questionnaire has been validated and is ready for use you will be able to use the measure as part of your routine evaluation methods. The measure will be sufficient to provide feedback regarding perceived QoL status to the participant and also as part of evaluation feedback to stakeholders such as PCTs. It is expected further work will be undertaken to validate the measure so that it will be able to detect change in QoL over time (i.e., validated for use pre-post intervention). There may also be provision to translate the measure into other languages as required.

The questionnaire you have been provided with for testing (enclosed) will not reflect the content and format of the final measure and therefore will not be validated at this stage. The final measure will be much shorter in length. For copies of the final validated measure and other evaluation tools you should contact:

Charlotte Hilton:
Email: charlotte.hilton@ntu.ac.uk
Telephone: 07939247921

Toni Minniti:
Email: antoinette.minniti@ntu.ac.uk
Telephone: 0115 88483918

Appendix 18 Administration Protocol for Recruitment Sites

Validation of an Exercise Referral Sensitive Quality of Life Measure

The following is a list of procedures that have been developed to help you administer the questionnaire:

1) For the purposes of this research study the following participants are eligible to complete the questionnaire:

- Anyone who has started their programme of exercise. However, **they must have participated in the exercise they are undertaking on ONE OR MORE occasions.**
- Anyone who has completed their initial referral for exercise and are maintaining their exercise programme (up to 12 months).

2) Please ensure that each person completing the questionnaire is given a consent form (copies enclosed). Please return all consent forms to Nottingham Trent University along with each corresponding questionnaire. There will be one consent form per two questionnaires if you are administering on two separate occasions.*

3) Please ensure that each person completing the questionnaire(s) is given **both a short and a long version of the information sheet** (copies enclosed) and has time to read the **short version** before completing the questionnaire. They may take the longer version away with them to read in their own time.

4) Please ensure that respondents have answered **all** the questions and that no responses are missing.

5) *If participants are completing two copies of the questionnaire, please administer the questionnaire for the second time **SEVEN** days apart from their initial completion. You should return both copies to Nottingham Trent University at the address provided.

Please return completed questionnaires to:
Charlotte Hilton
Nottingham Trent University
Clifton Campus
Room 204
Erasmus Darwin
NG11 8NS
Email: charlotte.hilton@ntu.ac.uk
Telephone 07939247921

Appendix 19 The Final Scale: The ER-QLS

The Exercise Referral Quality of Life Scale (ER-QLS)

Participant Name_____ Code_office use____

Date of birth_____

Gender (please tick) Male ☐ Female ☐

Primary Reason(s) for Referral_____
_____

Date Questionnaire was Administered_____/_____/_____

Stage of Referral when Questionnaire was Administered (please tick):

Exercise Programme Start ☐
Mid Exercise Programme ☐
Post Exercise Programme ☐

Scheme Name/Location_____

Ethnicity (please tick): White ☐ Mixed ☐
Indian ☐ Pakistani ☐ Bangladeshi ☐
Other Asian ☐ Black Caribbean ☐
Black African ☐ Other Black ☐ Chinese ☐
Other Ethnic Group ☐

Please answer all the questions you are asked. If you are unsure about which response to give to a question, **please choose the ONE** that seems most appropriate. This can often be your first response. At some points during the questionnaire response options will change so we ask that you answer carefully.

We ask that you think about your **life in the last two weeks** when responding to the questions that will follow. Please **circle** one number only.

	1 Very Poor	**2 Poor**	**3 Neither good nor poor**	**4 Good**	**5 Very good**
1) How would you rate your outlook on life?	1	2	3	4	5
2) How would you rate your level of self-esteem?	1	2	3	4	5
	1 Very Poor	**2 Poor**	**3 Neither good nor poor**	**4 Good**	**5 Very good**
3) How would you rate your physical health?	1	2	3	4	5
4) How would you rate your sense of well-being?	1	2	3	4	5
5) How would you rate the quality of your sleep?	1	2	3	4	5
6) How would you rate your memory function?	1	2	3	4	5
7) How would you rate your ability to manage stress?	1	2	3	4	5
8) How would you rate your ability to manage your mood?	1	2	3	4	5
9) How would you rate your ability to manage your weight?	1	2	3	4	5
10) How would you rate your current life expectancy?	1	2	3	4	5

11) How would you rate your overall quality of life?	1	2	3	4	5

	1 Never	**2 Seldom**	**3 Quite often**	**4 Very often**	**5 Always**
12) How much does any injury you may have prevent you from being physically active?	1	2	3	4	5
13) How often do you experience physical pain?	1	2	3	4	5
14) How often do you feel you have to manage any symptoms of pain?	1	2	3	4	5
15) How often do you feel you have to manage the symptoms of any illness you have?	1	2	3	4	5
	1 Not at all	**2 Not much**	**3 Moderately**	**4 A great deal**	**5 Completely**
16) How much do you currently enjoy physical activity?	1	2	3	4	5
17) How confident are you using gym equipment?	1	2	3	4	5

18) How confident are you that you can exercise in a leisure centre with minimum support?	1	2	3	4	5
	1 Not at all	**2 Not much**	**3 Moderately**	**4 A great deal**	**5 Completely**
19) How confident are you in your own physical ability?	1	2	3	4	5
20) How confident are you to join in activities with family and friends?	1	2	3	4	5
21) In general, how much opportunity do you have to make time for yourself?	1	2	3	4	5
22) How often do you feel you have achieved something in terms of your physical fitness?	1	2	3	4	5

Thank you for completing this questionnaire

Appendix 20 SPSS Syntax File for Scoring the ER-QLS

To calculate the total score of the ER-QLS electronically, copy and paste the following into a blank syntax file in SPSS:

```
RECODE Q12R Q13R Q14R Q15R (1=5) (2=4) (3=3) (4=2) (5=1).
EXECUTE.

COMPUTE ERQLS_TOTAL=Q1 + Q2 + Q3 + Q4 + Q5 = Q6 + Q7 + Q8 + Q9 +
Q9 + Q10 + Q11 + Q12R + Q13R + Q14R + Q15R + Q16 + Q17 + Q18 + Q19 +
Q20 + Q21 + Q22.
EXECUTE.
```

To calculate the total scores from each of the three domains, copy and paste the following into a blank syntax file in SPSS:

```
COMPUTE MPW_TOTAL= Q1 + Q2 + Q3 + Q4 + Q5 + Q6 + Q7 + Q8 + Q9 +
Q10 + Q11.

COMPUTE IPI_TOTAL= Q12 + Q13 + Q14 + Q15.

COMPUTE PAF_TOTAL= Q16 + Q17 + Q18 + Q19 + Q20 + Q21 + Q22.
```