

The Usefulness of ‘Think-Aloud’ for Evaluating Questionnaires in use in the Health Domain

A thesis submitted to the University of Manchester for the degree of Doctor in
Clinical Psychology (ClinPsyD) in the Faculty of Medical and Human Sciences

2014

Anna C. E. Phillips

Section for Clinical and Health Psychology

School for Psychological Sciences

Table of Contents

Table of Contents

Table of Contents	2
List of Tables and Figures	5
Abstract	6
Declaration	7
Copyright and Ownership	7
Acknowledgements	8
Paper 1: Systematic Review	9
Abstract	10
Introduction	11
Method	12
Quality Appraisal	14
Data Abstraction and Synthesis	16
Results	17
Study Quality	21
Overview of Studies	21
Aims of the Studies	22
Justification of Think-aloud Methodology	22
Populations	24
Samples	24
Sample Size	25
Education	25
Language	26
Methodology	26
Analysis	27
Evaluation of the Usefulness of Think-aloud Techniques in Addressing the Aims of the Study	27
Discussion	30
Strengths and limitations of the current review	32
Conclusions	33
References	34
Paper 2: Research Paper	40

Introduction	42
Method	46
Participants	46
Measures.....	46
Design and Procedure	46
Analysis	47
Results.....	49
Sample characteristics	49
Distribution of problems across the questions	52
Nature of the problems identified	52
Participants' Opinion of the ERQ.	56
Rationale for selecting the middle response-option	56
Spontaneous emotion-related comments.....	57
Discussion.....	58
Distribution and nature of problems with the ERQ.....	59
Interpretation of the middle response-option	60
Strengths and limitations of the current study.....	61
References	63
Paper 3: Critical Appraisal	67
Introduction	68
The Original Study.....	70
Transition to the Current Thesis	72
Learning from the Process	72
Methodological Challenges and Development of the Current Thesis	73
Paper 2	75
The think-aloud method	75
The Sample.....	76
Paper 2: Conclusion	77
Paper 1	78
Process Issues	78
Mixed-Methodology	79
Key Findings	80
Overall Conclusion	81
References	82

Appendices.....	88
Appendix A: British Journal of Health Psychology: Author Guidelines.....	88
Appendix B: Confirmation of Ethical Approval for the original study from NRES Committee North West - Liverpool East.....	94
Appendix C: Confirmation of Ethical Approval for the Substantial Amendment from NRES Committee North West - Liverpool East.....	100
Appendix D: R&D Letter of Approval of Substantial Amendment.....	105
Appendix E: The Emotion Regulation Questionnaire (ERQ; Gross & John, 2003)	108
Appendix F: Demographic Questionnaire.....	110
Appendix G: Participant Information Sheet.....	113
Appendix H: Participant Consent Form.....	118
Appendix I: CASP Quality Appraisal Tool	121

Total word count: 19,646

List of Tables and Figures

Figure 1- Literature Search: Procedure	14
Table 1- Summary of the Studies Included in the Review	16
Table 2- Number of Participants Demonstrating Each Demographic and Clinical Characteristic.....	46
Table 3- Number of Participants Coded as Having Any Type of Problem for Each Question	48
Table 4- Frequency and Exemplar Problems	49
Table 5- Frequency and Nature of Problems for Each Question.....	50
Table 6- Frequency and Type of Problem for Cognitive Reappraisal and Emotional Suppression Questions	52
Table 7- Coded Rationales for Selection of the Middle Response-Option	53

Abstract

The Usefulness of 'Think-Aloud' for Evaluating Questionnaires in use in the Health Domain.

A thesis submitted to the University of Manchester for the degree of Doctor in Clinical Psychology (ClinPsyD). Anna C. E. Phillips, 2014.

Self-report questionnaires are frequently used in health fields; however, subjective interpretation is often ignored. One way of assessing this is using techniques derived from cognitive interviewing. Of these, 'think-aloud', in which respondents speak their thoughts aloud as they complete a questionnaire, is the original paradigm. The thesis focusses on the use of 'think-aloud' methodology in the evaluation of questionnaires already in use in the health domain. The current thesis has been prepared in the format of scientific papers.

Paper 1 is a systematic review (23 studies) of the appropriateness and usefulness of think-aloud techniques for evaluating health-related questionnaires. A descriptive account is provided of the aims of the studies reviewed; the justification for using think-aloud; populations studied; and methodology; an evaluative account depicts the usefulness of the think-aloud method in addressing researchers' aims. Think-aloud was successfully used to address researchers' aims and was effective at elucidating problems with questionnaires. Theoretical and clinical implications are discussed, and recommendations made for future research.

Paper 2 is a cross-sectional observational study using think-aloud methods to examine the way in which people with End Stage Renal Disease (N=25) interpret and respond to the Emotion Regulation Questionnaire (ERQ). All questions were found to be problematic to some extent and reappraisal questions yielded the most problems. A tendency to deny or minimise negative emotions and present a positive self-image was also noted. Implications are discussed for use of the ERQ and replication with further samples suggested.

Paper 3 is a critical appraisal of the above papers and provides personal reflections on the research process as a whole. The current thesis was a transition from a different study; amended due to time constraints. The journey is also outlined from this original study to the present thesis. Strengths and limitations are considered as well as areas for improvement and future research.

Declaration

No portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Copyright and Ownership

The author of this thesis (including any appendices and/or schedules to this thesis) owns certain copyright or related rights in it (the “Copyright”) and s/he has given The University of Manchester certain rights to use such Copyright, including for administrative purposes.

Copies of this thesis, either in full or in extracts and whether in hard or electronic copy, may be made only in accordance with the Copyright, Designs and Patents Act 1988 (as amended) and regulations issued under it or, where appropriate, in accordance with licensing agreements which the University has from time to time. This page must form part of any such copies made.

The ownership of certain Copyright, patents, designs, trade marks and other intellectual property (the “Intellectual Property”) and any reproductions of copyright works in the thesis, for example graphs and tables (“Reproductions”), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property and/or Reproductions.

Further information on the conditions under which disclosure, publication and commercialisation of this thesis, the Copyright and any Intellectual Property and/or Reproductions described in it may take place is available in the University IP Policy (see <http://www.campus.manchester.ac.uk/medialibrary/policies/intellectual-property.pdf>), in any relevant Thesis restriction declarations deposited in the University Library, The University Library’s regulations (see <http://www.manchester.ac.uk/library/aboutus/regulations>) and in The University’s policy on presentation of Theses

Acknowledgements

I would like to acknowledge all who have contributed to the completion of this thesis. Special thanks to the patients who took part and the staff of the Renal Units; in particular, Sandip Mitra, Jon Courthold and Rincy Sajith. I am exceptionally grateful to my supervisors Professor Alison Wearden and Dr Dougal Hare for their time, contribution and encouragement throughout this process.

I would also like to thank my wonderful colleagues for the camaraderie over the past three years, and of course, my family, friends and boyfriend for their unfailing love, patience and support.

Paper 1: Systematic Review

The Usefulness of ‘Think-Aloud’ for Evaluating Questionnaires in the Health Domain: A Systematic Review.

The following paper has been prepared with submission to the ‘British Journal of Health Psychology’ in mind.

The guidelines for authors can be found in appendix A

Word Count: 7,778

The Usefulness of ‘Think-Aloud’ for Evaluating Questionnaires in the Health Domain: A Systematic Review.

Abstract

Purpose. Measurement of health-related outcomes informs understanding of health, delivery of interventions and health-care planning, and is frequently undertaken using survey questionnaires. It is vital that researchers can be sure that these instruments measure what they purport to measure. One way of determining this is using techniques derived from cognitive interviewing. Of these, ‘think-aloud’, in which respondents speak their thoughts aloud as they complete a questionnaire, is the original paradigm. Recently, think-aloud has been increasingly used to evaluate questionnaires in the health domain. The present study is a systematic review (23 studies) of the appropriateness and usefulness of think-aloud techniques for evaluating health-related questionnaires.

Method. A systematic database search was conducted. Papers were included if they used concurrent think-aloud methodology to evaluate questionnaires currently in use in the health domain.

Results. The review presents a descriptive account of the aims of the studies reviewed; the justification for using think-aloud; populations studied; and methodology. Think-aloud was successfully used to address researchers’ aims and was effective at elucidating problems with questionnaires.

Conclusions. Frequently-used questionnaires are not consistently understood or completed as researchers intended. Results suggest that further research into the validity of health-related questionnaires is warranted; moreover, this can be effected successfully using think-aloud techniques.

Introduction

Accurate measurement of health-related outcomes is central to our understanding of health and illness at both an individual and epidemiological level. Health data informs the development and evaluation of health interventions, as well as health policy and public accounting. McDowell (2006) asserts that subjective reports are the most economically viable evaluative tool to assess health and to provide unique insight into the subjective experience of the respondent. Critically, the data gathered from self-report is only useful to the extent that people make sense of the questions in the manner intended; that is, to the extent to which the scale is measuring the variables it purports to assess.

Cognitive interviewing (also known as think-aloud interviewing or verbal protocols) has enabled researchers to test the assumption, inherent in standardised surveys, that respondents are 1) universally able to understand the questions being asked of them; 2) that this understanding is the understanding intended by the researcher; and 3) that they are subsequently able, and willing, to answer such questions (Collins, 2003). The method has been used to elucidate the way in which respondents understand and interpret questions, and to identify potential problems with questionnaires (Drennan, 2003); including health-related questionnaires (e.g. Horwood, Sutton & Coast, 2013; Heesch, van Uffelen, Hill & Brown, 2010; McCorry, Scullion, McMurray, Houghton & Dempster, 2013).

There are two main techniques in cognitive interviewing. The first and original paradigm is concurrent think-aloud (Ericsson and Simon, 1980, expanded in 1993). In this, respondents are instructed to verbalise their thoughts as they read and complete a questionnaire. The second is verbal probing (for example, Converse and Presser, 1986), in which the interviewer questions the subject either concurrently or retrospectively regarding, for example, their understanding, opinions, or judgements. Whereas the latter is interviewer-led, enabling researchers to guide the dialogue towards areas of interest (Willis, 1991), the former is directed by respondents, and considered more akin to ‘real-life’ survey completion (Conrad, Blair, and Tracy, 2000) (See Beatty & Willis, 2007 for an assessment of the two paradigms). Beatty and Willis (2007) note that whilst cognitive interviewing has diversified from original think-aloud procedures, virtually all descriptions still include concurrent think-aloud as one possible component (see DeMaio and Landreth 2004; Willis

2005), and some researchers (e.g., Conrad, Blair, and Tracy 2000) have continued to favour the approach. The focus of this review will be on the use of concurrent think-aloud with questionnaires in the health domain.

To date, think-aloud protocols have been used extensively in the pre-testing of health questionnaires; that is, in the stages of design, development, adaptation or translation of health surveys (see Drennan, 2003 for a review). It cannot be assumed however, that all questionnaires have been assessed in this way. Furthermore, one cannot necessarily generalise think-aloud data from one population to another with varying clinical, age, socioeconomic, cultural or gender profiles. For these reasons, think-aloud interviewing can be a valuable tool in the evaluation of questionnaires currently in use in the field (e.g. Darker & French, 2009; van Oort, Schroder & French, 2011).

The current review aims to identify the extent to which think-aloud methods have been used with questionnaires already in use in the health domain; to identify the aims of their use; and to provide a preliminary evaluation of the usefulness of think-aloud methods in meeting these aims. Beatty and Willis (2007) highlight a lack of consensus within the cognitive interviewing literature upon methodological issues such as sample size and participants. Accordingly, the review will describe the populations, sample sizes and approaches within the papers identified.

Method

The electronic databases CINAHL-Plus, Medline and PsycINFO were searched using the following search terms: “Think*-Aloud” or “Think* Aloud” or “Cognitive Interview*” and Questionnaire or Survey or Measure* or Interview* or Scale*. Manual searches of the reference lists of identified papers followed. The current search was conducted in January 2014.

Inclusion criteria for selecting papers were:

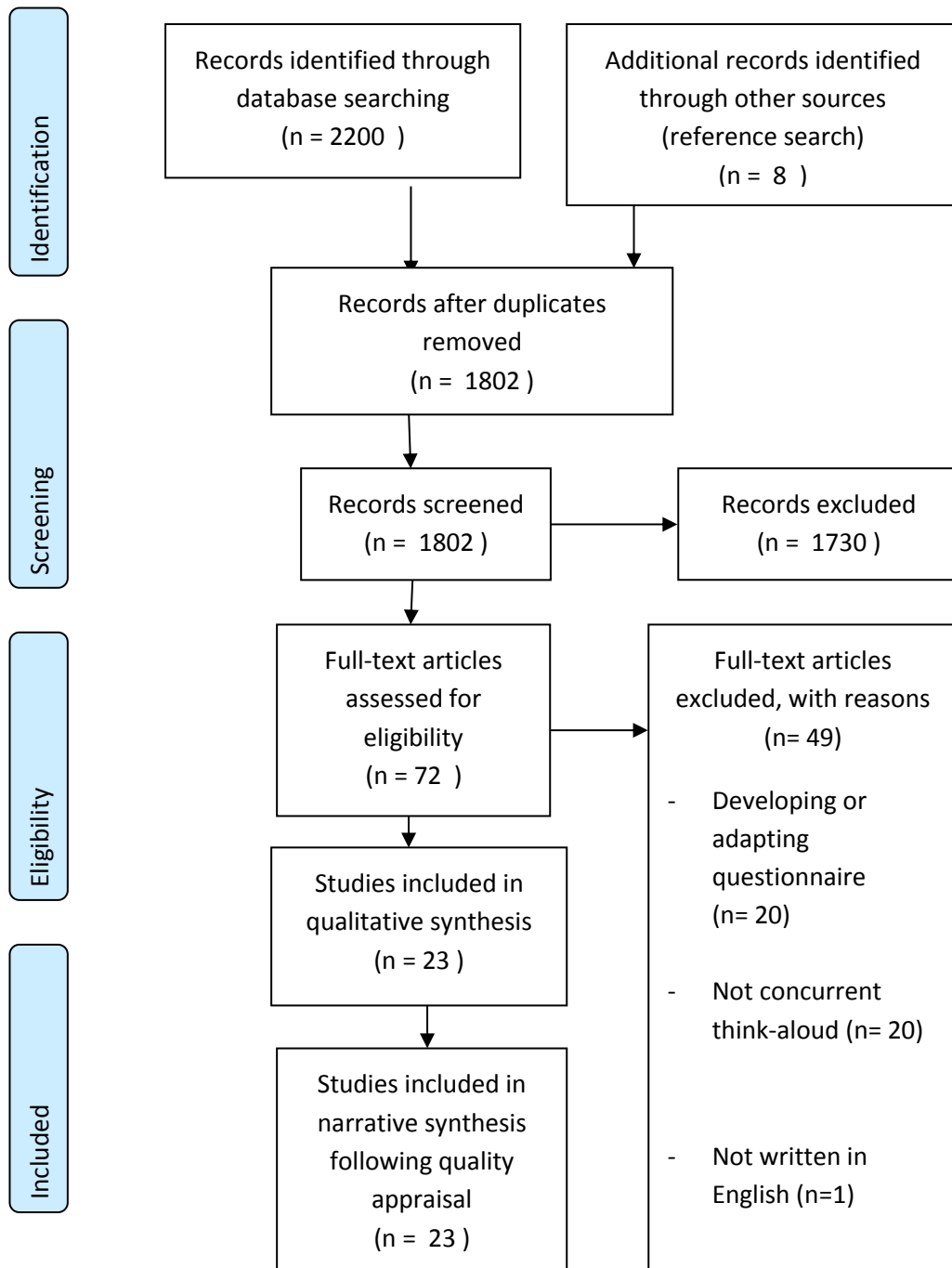
- Use of concurrent think-aloud methodology
- Think-aloud used to evaluate health-related questionnaires
- Questionnaires evaluated are currently in use in the health domain
- Published in the English language in peer-reviewed journals.
- Primary research studies

Exclusion criteria:

- Studies conducted with staff groups only
- Think-aloud used in the process of pre-testing (i.e. developing or adapting questionnaires).

Figure 1 details the phases of the systematic literature search. The database searches produced 2200 records. The reference manager Endnote was used to remove duplicates resulting in 1794 articles. These were screened and 1730 excluded as they did not meet inclusion criteria. The remaining 64 papers were considered in detail. 22 studies met inclusion criteria. A reference search of these articles identified 8 further potentially relevant papers. Upon further examination only one of these met criteria for inclusion in the review. The likely reason that this paper was not identified from the database search was that the title and keywords did not include the term 'think-aloud'. All searches revealed that no similar systematic review had previously been published to date.

Figure 1. Literature Search: Procedure



Quality Appraisal

The methodological quality of the papers was evaluated in order to guide the interpretation of findings. A quality assessment tool was used to ensure standardised assessment (Centre for Reviews and Dissemination, 2009). In this case the Critical Appraisal Skills Programme tool for qualitative research (CASP, 2001) was used. This is a 10 item checklist, each with guidance, or ‘hints’, on what to consider when appraising each item (see appendix I). The items appraise the clarity, sufficiency and appropriateness of: the aims of the research; the design; methodology; recruitment strategy; data collection; and analyses. It also evaluates whether studies have taken into account ethical issues and the relationship between the researcher and participants. The CASP Qualitative Tool was chosen as the review centres around papers using think-aloud, an inherently qualitative method. The first author judged the extent to which the papers met each item on the quality checklist, using the following criteria: if there was sufficient evidence to demonstrate the item, they were scored a 2; if the item was partially demonstrated, they were scored a 1; and if there was no evidence of the item, they were scored a 0. For example, for item 3, ‘Was the research design appropriate to address the aims of the research?’, the researcher considered the ‘hints’ provided in the checklist: ‘has the researcher justified the research design? Have they discussed how they decided which method to use?’. If the design was judged as appropriate to address the aims, and the paper discussed and justified their choice of method, then the paper was scored a 2. If the design was judged appropriate, however, there was insufficient discussion or justification of the method then the paper was scored a 1. If none of these criteria were met, then the paper would be scored a 0.

To address issues of subjectivity in the use of this tool, a sample of the papers was independently evaluated by a second researcher, using the above criteria. The two researchers agreed on 9 out of the 10 papers sampled; giving a measure of agreement at 90%. The remaining paper was agreed following discussion. Papers were included in the current review based on best evidence synthesis (BES; Popay, Roberts, Sowden, Petticrew, Arai, Britten et al., 2006). In BES, studies must meet minimum standards of methodological adequacy and relevance to qualify for inclusion in the review. To this end, papers were required to meet a minimum quality standard. In this case, scores from the 10 items were totalled for each paper. It was decided that

papers must meet a score of 10 or above on the CASP quality appraisal tool for inclusion in the review (see 'Results', 'study quality' for details). All items on the CASP were equally 'weighted' in terms of their contribution to the overall quality score.

In the current study, quality ratings were conducted primarily to determine whether studies were of sufficient quality to be included in the review. At synthesis, studies were not weighted in terms of whether they were moderate, high, or very high quality. All papers were considered equal in the narrative synthesis.

Data Abstraction and Synthesis

A narrative synthesis was conducted (Centre for Reviews and Dissemination, 2009). This was conducted systematically by the first author, based upon pre-existing research questions (e.g. the justifications given for using think-aloud; the aims reported). For each research question in turn, the papers were examined for evidence. For example, in exploring the justifications given for think-aloud, each paper was scoured for justifications and all justifications noted. Once justifications were obtained for all papers, these were considered together. All similar justifications were grouped and the number of papers reporting these justifications was reported in the results. Justifications identified by as few as one paper were reported. The majority of the research questions were descriptive (e.g. justifications and aims); one ('evaluation of the usefulness of the think-aloud technique in addressing the aims of the study') was interpretative.

Results

Table 1

Summary of the Studies Included in the Review

Authors, year & location	Aim/ Objective	N	Sample characteristics	Measure	Method	Analysis	Key Findings	Quality Rating Score
1. Al-Janabi et al. (2013) UK	Examine the feasibility of self-reporting capability	34	Diverse sample (age, sex, socioeconomic status, health status on EuroQoL-VAS) 16 male, 18 female	ICECAP-A	Concurrent think-aloud & interview	Content analysis	Problems on fewer than 10% of the items. Overall participants responded to capability questions in the intended manner.	18
2. Bode & Jansen (2013) The Netherlands	Examine contextual validity of the PEAS	30	Patients with arthritis. 22 women, 8 men; aged 43-84 yrs (mean=64 yrs, SD = 11)	PEAS (physical decline subscale)	TSTI	Qualitative analysis	Some problems were identified due to faulty item formulation. Overlap in meaning between the experience of being chronically ill and the experience of aging.	19
3. Brown, et al. (2009) UK	Investigate construct validity of the ASI-R	16	Patients with a primary diagnosis of anxiety 10 female, 6 male; mean age 37.5 yrs	ASI-R	Concurrent think-aloud	Content analysis	The ASI-R does not measure beliefs as it purports to measure.	16
4. Darker & French (2009) UK	Understand the processes of TPB completion	45	Public & University staff 26 females, 19 males, Mean age 32.6 yrs (SD = 11.7)	TPB	Concurrent think aloud	Content analysis, T-tests & Chi Squared	16 problems with 52 questions. Problems were associated with increased endorsement of the middle response-option. Normative and intention questions were particularly problematic.	20
5. French et al. (2007) UK	Identify problems with TPB questionnaire	Study 1= 6 Study 2 =13	1. All have risk factors for poor health. 30–50 yrs; 2 males, 4 female. 2. University students, 18–26 yrs; 7 females, 6 males.	TPB	Concurrent think aloud	Content analyses & Chi Squared	Most people had no problems with most questions. However, there were problems common to both studies. Normative influence questions were particularly problematic.	17
6. French & Hevey	Identify thoughts in	40	University students. 20 male, 20 female; aged 18 –	3 items to assess risk	Concurrent think aloud	Content analysis &	Most common thoughts concerned exposure to the sun & features such as	19

(2008) UK	measures of unrealistic optimism		24 yrs		of skin cancer		t-tests	skin colouring. Different thoughts were noted for direct & indirect questions. Numerical probabilities rarely considered.	
7. Galasinski (2008) UK	Examine the discourses in completing the BDI	52	Healthy men and women with no previous psychiatric care. Range of: ages; education; & domicile (not reported)	BDI (Polish version)	Concurrent think-aloud	Qualitative analysis	Most participants found the items problematic, unacceptable. These were instances where the experience of the informant could not be framed by the BDI.	15	
8. Heesch et al. (2010) Australia	Document understanding of the IPAQ	41	Sufficiently mobile older adults 19 men, 22 women; 65-89 yrs (mean = 72.9)	IPAQ	Concurrent think-aloud & probes	Qualitative analysis	Most difficulties during the primary task stages. Most problems with the moderate-intensity PA & walking questions.	18	
9. Horwood et al. (2013) UK	Identify problems with the ICECAP-O	20	Patients with osteoarthritis of the knee or hip. 14 female, 6 male 48-87 yrs	ICECAP-O capability measure	Concurrent think-aloud	Content analysis & thematic analysis	Problems in 7 % of question segments. Majority were comprehension problems; no retrieval problems were identified. Identified differential number of problem with each of ICECAP-O's concepts.	20	
10. Kaklamanou et al. (2013) UK	Identify difficulties completing Scale & steps to improve	43	University students & staff 32 female, 11 male; 20-63 yrs (mean=34.98, SD = 13.94)	Compensatory Health Beliefs Scale	Concurrent think aloud	Thematic analysis	Evidence of internal reliability & face validity. Most responses not problematic, yet, all participants had problems & all items were problematic; some more than others. Conceptual ambiguities identified.	19	
11. Lahaut, et al. (2003) The Netherlands	Compare data quality of 2 weekly recall & 2 Quantity Frequency alcohol scales	17; 11; 12; 12	Variation in educational level and drinking pattern. (Demographics not stated)	WR1 & WR2; QF1&QF2	Concurrent think-aloud & probes	Content analysis	Problems on all measures. Most on WR1&WR2 due to recall difficulties. WR1 most problematic due to instructions. Serious problem with QR1&QR2 in calculating average consumption leading to over reporting.	15	
12. Mallinson (2002) UK	Explore interpretation of the SF-36	56	Referred for Physiotherapy/ Occupational Therapy. 12 men, 44 women; aged 65-89 yrs (mean=77)	SF-36	Concurrent think-aloud	Qualitative analysis	Problems identified. Difficult to interpret answers as people have different meanings/ intentions in selecting a response option.	18	
13. McCorry, et al. (2013)	Access cognitions whilst	36	Patients with diabetes 18 male, 18 female mean age = 64.77 years (SD	IPQ-R	Concurrent think-aloud	Content analysis	Most items not problematic. Problems with the concept of 'cure' and 'symptoms' & with the negative phrasing. Similar	17	

UK	completing IPQ-R,	10.00)					scores may reflect very different cognitions about diabetes.	
14. Mulcahey, et al. (2011) USA	Evaluate the items on the SAQ.	76	31 youths with scoliosis & 45 typically developing (8-16 yrs; mean= 13)	SAQ	Concurrent think-aloud; probes & interview	Content analysis & Nonparametric statistics	Reading & comprehension problems & problems understanding illustrations for all items. Results consistent regardless of age or diagnoses.	18
15. Pool, et al. (2010) The Netherlands	Identify problems with the PCCL & interpret questions.	13	Participants with non-specific sub-acute neck pain (demographics not reported)	PCCL	Three Step Test Interview (TSTI)	Qualitative analyses	43% of items were problematic. Different types of problems elucidated. Response category sometimes mis-used.	18
16. Taminiau-Bloem et al. (2010) The Netherlands	Examine consistency in QoL appraisal	50	Cancer patients. 24 male, 26 female; aged 30-80. Selected for mix of: tumor site & length of radiotherapy treatment	EORTC QLQ-C30	TSTI	Qualitative analysis	In 94% of comparisons of responses, the content of at least one cognitive component changed over time. No patterns of (dis)similarity discerned.	19
17. Taminiau-Bloem et al. (2011) The Netherlands	Study Assumptions of transition questions on the EORTC QLQ-C30	25	Cancer patients. 12 male, 13 female; aged 30-80. Selected for mix of: tumor site & length of radiotherapy treatment	EORTC QLQ-C30	TSTI	Content analysis	In 112/164 responses patients compared current and prior functioning. However, 104 of these used a variety of time frames. In 79 responses, the time frame &/or description of prior functioning differed from those at pretest.	20
18. van Oort et al. (2011) The Netherlands	Assess the content validity of the Brief IPQ	Study 1=6 Study 2= 11	1. Patients waiting for elective colon/AAA surgery, doing pre-operative exercise. 4 male, 2 female (aged 54–78 yrs). 2. Musculoskeletal problems. 8 female, 3 male; aged 18-87.	Dutch version of the Brief IPQ	Concurrent think aloud	Content analysis	88 problems identified. The pattern was similar across both samples. Identity, personal control, illness coherence, and causal attribution questions yielded most difficulties.	18
19. Watanabe, et al. (2008) Canada	Gather validity evidence for ESAS	20	20 Advanced cancer Patients. 10men, 10 women; aged 41–74 yrs (Median=56)	ESAS	Concurrent think aloud & Interview	Content analysis	Symptom ratings influenced a number of different factors. Symptom interpretation and numerical rating assignments varied. Difficult terminology identified.	19

20. Westerman, et al. (2007) The Netherlands	Examine responses to question 'were you tired?' at different time points	23	Small-cell lung cancer patients. 12 had limited disease: 3 male and 9 female; aged 42–69 (mean=55); 11 had extended disease: 8 males, 13 females; aged 39-72 (mean=64)	EORTC QLQ-C30 question 'were you tired?'	TSTI	Qualitative analyses	15/23 discrepancies between answers & spontaneous reports of fatigue. These rated fatigue as minimal, with varying explanations. In all, response shift & a change in perspective were found. Spontaneous reports suggested protective & assertive behaviour & fighting stigma.	18
21. Westerman, et al. (2008) The Netherlands	Investigate interpretation & response to EORTC-QLQ-C30	23	Small-cell lung cancer patients. 12 had limited disease: 3 male and 9 female; aged 42–69 (mean=55); 11 had extended disease: 8 males, 13 females; aged 39-72 (mean=64)	EORTC-QLQ-C30 (GH/QOL; PF; & RF scales)	TSTI	Qualitative analyses	Various response strategies used to answer questions about limitations in functioning, reducing the accuracy of the scale. Limitations underestimated due to overly-literal interpretations; guessing; & ignoring/excluding challenging activities.	18
22. Willis, et al. (1991) US	Examine problems in survey questions	Study 1=24 Study 2= 19	1. People who used an assistive device 2. Public (Demographics not reported)	1.Assisted Devices Questionnaire; 2.Radon questionnaire	Concurrent think aloud & probes	Mixed methods	39 problems found in the assistive devices questionnaire, 12 were structural in nature, and 27 were cognitive. For the radon questionnaire, all seven problems were predominantly cognitive.	14
23. Wylde, et al. (2012) UK	Explore issues completing pain questionnaire.	20	Moderate–extreme pain in replaced joint (10 knee, 10 hip). 10 female, 10 male; Mean age=69 yrs,	CPQ	Concurrent think aloud	Thematic analysis	Problems identified with CPQ. Influenced by:1) challenges with the question wording/response options; 2) fluctuating pain & functional limitations; co-morbid pain & problems & adjustment to pain.	17

Study Quality

Quality ratings are listed in table 1. For the current review, scores of 1-10 were considered 'low' quality (criteria for exclusion from the review); 11-15 'moderate'; 16-18 'high'; and 19-20 'very high' quality. No studies were excluded on the basis of insufficient quality. 12 studies demonstrated 'high' quality and 8 showed 'very high' quality. Deficits in quality were due to low scores upon criteria '6' of the CASP tool: 'Has the relationship between researcher and participants been adequately considered?'. In many papers, there was minimal or no consideration of the role of the interviewer and potential bias upon the data collected (5, 1-3, 6, 8, 10-16, 18-23). Further deficits were the result of not evidencing ethical considerations (3, 5, 7, 20-22). The results of three studies must be interpreted with caution due to 'moderate' quality ratings (7, 11, 22). These studies demonstrated poorer quality in the reporting of methodology (processes of recruitment, data collection and analysis) and findings. As indicated in 'quality appraisal' above, in the narrative synthesis, studies were not weighted in terms of their quality rating score. All papers were treated equally at synthesis; however, it is likely that assigning a greater weighting to higher quality papers would have influenced the results at this stage. In this case, greater emphasis would have been placed on the results of the 8 studies with 'very high' and 12 studies with 'high' quality compared to the results of the three with moderate quality ratings. It should be noted that whilst the CASP was employed in the current study, the use of a different quality appraisal tool might also have elicited different results. Furthermore, scores of low, moderate, high and very high quality were developed idiosyncratically and the use of alternative 'cut off' scores may have resulted in papers receiving a different rating. More stringent judgements of low quality (such as a score of 15) would have led to the exclusion of three studies in this review. Also, if certain items were weighted more heavily in appraising study quality (for instance, the appropriateness of the methodology, data collection and analysis), then studies may have been excluded due to low quality. Whilst this may have added to inferences regarding methodological rigour and quality of the studies reviewed, it also risks excluding important findings, presented in an alternative format.

Overview of Studies

In accordance with inclusion criteria, all studies employed a concurrent think-aloud (qualitative, quantitative or mixed-methods) design (N= 23). Fourteen studies sampled participants from a physical health population; one from a mental health

population; and nine from a non-clinical population. The review is synthesised under the following headings: aims of the studies; justification in using think-aloud methodology; populations; methodology; and preliminary evaluation of the usefulness of think-aloud in meeting aims.

Aims of the Studies

The papers in this review reported two predominant aims in using the think-aloud methodology. The first (reported by 11 studies) was to assess the utility of a questionnaire. The majority of these sought to explore the extent and nature of problems or issues which people encounter in completing a questionnaire (5, 9-10, 15, 19, 22-23) and one examined whether people responded as intended to instructions (3). Studies explicitly sought to assess content validity (18); construct validity (3); and the feasibility of self-reporting the variable of interest (1). One study used think-aloud to compare the effectiveness of two questionnaires to determine alcohol consumption (11) and two further sought to inform questionnaire improvement (2, 10).

The second aim (reported by 13 studies) was to understand respondents' thought processes while they were completing questionnaires. In the main, this related to interpreting and responding to questionnaires (2, 4, 6-7, 12-15), and four studies specifically explored cognitions surrounding time-point (20-21) or transition judgements (16-17). This aim is differentiated from the first, to the extent that it focusses upon cognitive processes underpinning the task of questionnaire completion; as distinct from the examination of the validity or feasibility of a questionnaire.

Justification of Think-aloud Methodology

All studies discussed the nature of think-aloud methodology and all but two (16-17) provided at least a minimal justification for selecting the method. Reasons given were diverse; the predominant justification (14 studies) being that think-aloud has proven an effective tool in previous research. For instance, it was documented that think-aloud was used effectively to: assess questionnaires in the researcher's domain of interest (1, 5-6, 9); to test the contents, design, or validity of questionnaires (7, 9-10); to reliably observe and differentiate problems which people encounter in completing questionnaires (4-5, 10); and to investigate the processes involved in questionnaire completion and interpretation (5, 12, 18, 19). Think-aloud protocols

were also described as important means of investigating cognitive processing strategies during problem-solving, decision-making and judgement tasks (4) and multi-step tasks (1).

Second most common (8 studies) were statements to the effect that think-aloud provided insight into cognitions, including the content of short term memory and enables the exploration of how and why people arrive at their answers to questionnaire items (2, 7, 10, 13-14, 18). One study furthered this, stating that concurrent interviews have generated more information and insights into the decision-making process than retrospective methods (1).

A number of studies indicated that think-aloud was the most naturalistic method of enquiry, presenting the least distraction from the target task (4) and the least impact upon the usual experience of questionnaire completion (9). Indeed, a rationale was presented for capitalising upon a social need to comment upon and account for questionnaire choices (7) and for avoiding 'in-depth probing' to prevent significantly altering the interview dynamic (12).

Two justifications related to the usefulness of think-aloud in establishing whether questionnaires measure variables of interest in the manner intended by the researchers. Think-aloud was stated as useful in gathering validity evidence (8, 19), including face validity (9-10) and contextual validity (2), as well as testing accuracy of responding (11); for instance whether participants follow instructions (3) or understand questions as intended (8). Studies further noted that in uncovering potential sources of error; think-aloud techniques can be used to improve the credibility of data gathered from questionnaires (8, 14) and are recommended when developing new measures (14). As the earliest study in the field, Willis (1991) presents arguments for and against the potential utility of think-aloud and aimed to test the validity of these using think-aloud in the 'laboratory'.

Two studies employing the Three Step Test Interview (TSTI; Hak, van der Veer & Jansen, 2008; Jansen & Hak, 2005), advocated the observation-based component of think-aloud (20-21). They considered the TSTI sufficient to identify discrepancy between the 'theory' underlying questions and respondents' actual behaviour (2, 15).

In discussing their results, the majority of studies did not specifically reflect upon the usefulness of the think-aloud method. Those which did, concluded that their findings lent credence to conclusions regarding the efficacy of the method (4-5, 15, 22), and two reiterated earlier statements as to the ‘naturalistic’ benefits of the technique (9, 23). Willis (1991) highlighted however, the benefits of retrospective probing to further explore themes which emerge during the think-aloud interview (22).

Populations

The papers demonstrated that the think-aloud method can be used successfully with people of ranging ages, cognitive and physical capacity and level of education to explore questionnaires in the health domain.

Interestingly, none of the studies reflected upon the inherent acceptability of the think-aloud method to participants. For instance, whilst the technique yielded informative findings, nowhere is it noted whether participants experienced ‘thinking aloud’ in a positive or negative light. Of note, only two of the twenty three studies excluded participants upon the basis that they could not follow think-aloud protocols (10, 13). From this it can perhaps be inferred that the method ‘made sense’ to the people sampled.

Samples

Participants were mostly from physical health populations, including cancer patients (16-17, 19-21), people with, or at risk of, developing type-2 diabetes (13 and 5 respectively) and patients with arthritis (2, 9). Only one study sampled a mental health population (3) and the remaining nine employed a non-clinical sample (1, 4-8, 10-11, 22). Think-aloud was conducted in a range of locations, including participants’ homes and hospital settings.

Think-aloud is a flexible approach to assess questionnaires in the health domain, appropriate for use with a wide range of ages (8-89 years). The effective use of think-aloud with older adults and children implies that the procedure does not present too great a cognitive load. Older adults may be more likely to suffer from cognitive or physical impairment, in addition to cultural differences in experiences of completing questionnaires, yet in the main, they responded well to the think-aloud method. Only one study reflected that older participants struggled more at implementation compared to younger subjects (15).

People with physical or mental health difficulties may experience reduced cognitive functioning as a result of pain, medication, fatigue, distress or cognitive impairment. The present review suggests that such cognitive difficulties do not preclude the successful use of think-aloud. This is important considering that health-related questionnaires are widely used with clinical populations. All non-clinical and the majority of clinical samples provided full data sets. Of note, some studies commented that their sample did not represent the most severely ill or complex clients (12, 16-17 19-21). In these cases, patients were too unwell or declined to participate. A number even died prior to completing participation. These patients were the only subjects who terminated participation early (16-17, 20-21), suggesting that ‘thinking aloud’ was acceptable to most participants. Importantly, a number of studies specifically excluded participants with cognitive impairment or dementia, or those who had a brain tumour or were being treated with brain irradiation (3, 9, 16-18, 20). Despite its efficacy with other clinical groups, this implies that researchers considered a reasonable level of level of cognitive functioning necessary to effectively explore questionnaires using this method.

Sample Size

Think-aloud yielded informative results with a wide range of sample sizes: the smallest being 6 (5, 18); the largest 76 (14). The mean number of participants was 28. Whilst studies employed small samples, some authors present this as a limitation to the generalizability of findings (5, 3, 18, 23) and recommend the use of larger samples in future studies. One reflected that with 20 participants they were unable to statistically compare quantitative data on the basis of sample characteristics. Conversely a sample size of 45 enabled another to meaningfully quantify problems encountered with questionnaires.

Education

The think-aloud protocol was used successfully with people from a range of educational backgrounds (7-9, 11). Authors noted the importance of sampling a range of educational levels to allow the generalisation of findings (5, 8, 23). Indeed, some observed differences in the way that more educated participants read, interpreted and responded to questionnaires compared to less educated individuals (5, 15).

Language

Think-aloud was used effectively in a number of languages, including English, Dutch and Polish. It is unclear, however, whether the method could be employed with individuals who are non-fluent in the native language of the questionnaire and some studies excluded individuals on the basis of non-fluency (3, 9, 16-17 19).

Methodology

Think-aloud was used to evaluate a variety of questionnaires in the health domain, including measures of health-related quality of life (HR-QoL) (16-17, 20-21); pain (15, 23); health beliefs or perceptions (10, 13-14, 18); symptomology (2, 19); capability (1, 9); perceived risk (6); anxiety (3); depression (7); and theory of planned behaviour questionnaires (4-5). It appears that the method can be applied to lengthy questionnaires (64 items [16-17, 20-21]), although to minimise participant burden, some studies reduced item numbers (8).

In the current review, papers were selected on their use of concurrent think-aloud protocols. Almost half, however, used additional qualitative techniques to explore questionnaire completion. Four used follow-up ‘probe’ questions (Collins 2003; Jobe and Mingay 1989) to elucidate processes occurring during concurrent think-aloud (8, 11, 14, 22). A further three employed retrospective interviews to examine any issues or opinions regarding the questionnaire (1, 14, 19). Notably, six employed think-aloud as part of the Three Stage Test Interview (2, 15, 16-17, 20-21). This consists of three phases: 1) concurrent thinking aloud; 2) a retrospective interview; 3) a semi-structured interview. It is important to consider that in reporting results, these studies do not differentiate findings from each method. As such, results reported constitute an amalgamation of qualitative methods. Future research is required to examine the relative contribution of each of these methods in assessing questionnaires in the health domain.

Of the studies which specified their instruction protocols, most based instructions given on Green & Gilhooly (1996) and addressed any queries prior to commencing the procedure (1, 3-6, 10, 13, 18). In line with recommendations for best practice (Green & Gilhooly, 1996), ten studies employed a ‘warm up’ procedure. Three (1, 16-17) used Willis’ (1994) exercise which involves participants imagining and counting the windows in their home whilst speaking aloud their thoughts and

observations. One used a demonstration task (8), and the remaining studies allowed participants to practice with similar questions.

Analysis

The majority of studies (1, 3-6, 9, 11, 13-14, 17-19) used a quantitative approach to tally the frequency of various interpretations or problems (Willis, 2005). Many chose qualitative methodology (2, 7-8, 10, 12, 15, 20-21, 23) to identify recurrent themes or patterns in the data (Willis, 2005). Three used thematic analysis (9, 10, 23); however, most did not specify the qualitative procedure used. The remaining studies (9, 16, 22) utilized a combination of approaches. Three studies sought to evaluate patterns in variables of interest through the use of parametric and non-parametric statistical procedures (4, 6, 14).

Evaluation of the Usefulness of Think-aloud Techniques in Addressing the Aims of the Study

Where possible the current review presents only the efficacy of the concurrent think-aloud method. However, as noted above, studies do not clearly differentiate findings elicited by think-aloud techniques from those obtained during additional qualitative procedures.

The evaluation begins with the first aim, of assessing the utility of questionnaires in the health domain. Of those studies aiming to explore content validity or the nature and extent of problems encountered, 'thinking aloud' identified problems with a significant number of questionnaire items. Some studies recognised difficulties with between 7% (9) and 57% (22) of items. Think-aloud also elucidated the nature of these problems, for instance challenges in comprehension (5, 9, 19, 22), interpretation (18, 22) and problems using the scale (10, 15, 19). The method further highlighted questions or concepts which proved more challenging than others (5, 10, 15, 18-19) and indicated that people will apparently answer survey questions that they may not understand (22). This is in line with evidence from experimental research showing that survey respondents try to answer questions no matter how difficult they find it (Tanur, 1992).

Think-aloud successfully examined whether people responded as intended to instructions (3). It was found that whilst, in the main, participants answered as instructed; in the 5% of cases they did not, this proportion was large enough to

distort psychometric analyses (Waller, 1989). In addition, the method enabled researchers to assess the construct validity of an anxiety questionnaire, leading them to conclude that the questionnaire measured concepts other than the variable it purported to measure (3). It also lent support to the feasibility of self-reporting the concept of ‘capability’ (1). Think-aloud proved useful in identifying issues with, and allowing the comparison of, two different questionnaires to assess alcohol consumption (11). A limitation was noted however; that is, whilst the technique detected misreporting of information, data did not indicate the extent to which this occurred, or whether misreporting was equal across questionnaires.

Lastly, in aiming to inform questionnaire improvement, the technique facilitated the identification of minor adaptations necessary to resolve problematic items (2, 10). It also elucidated problems which were more conceptual in nature and would not be so easily remedied (2). Interestingly, whilst only two papers explicitly stated improvement of the questionnaire as a preliminary objective, on the basis of results found using think-aloud, the majority of studies later made recommendations to improve questionnaires (2, 4-5, 8-10, 13-19, 22), or questioned the overall validity of the questionnaire for the intended purpose (3, 7, 14, 20-21, 23).

The second aim sought to understand cognitions involved in questionnaire completion. In the main, this related to interpreting and responding to questionnaires (2, 6-8, 12-15). To this end, all studies, bar one (6), presented their findings in terms of problems encountered by participants. Think-aloud again elucidated the nature of, and cognitions underlying, these problems and identified items which were more challenging than others. The method highlighted idiosyncratic responses to problematic items, such as reformulating a question or the context of a question to suit, or rejecting the item outright (7). One study found that on some problematic items, participants are significantly more likely to select the middle response option and considered this finding a validation of the think-aloud method (4); another noted that problems led to over or under-reporting of physical activity by older adults (8). Again the method highlighted that people will often persist in completing problematic items (7).

In accordance with aims, think-aloud shone a spotlight on personal meanings and interpretations of questionnaire items. It was noted that people have varied

understandings of certain concepts, for example, illness or treatment control (12, 13) and that meanings such as illness and aging may overlap. Also, similar scores were found to reflect different cognitions. For instance, the same score could reflect very different profiles of illness representation, dependent on whether the respondent reflected on their ability to cope with the consequences of illness before endorsing their response to these items. This has implications for interpretation by researcher and clinicians (12, 13).

The think-aloud technique further afforded insight into the way that subjects make QoL judgements, for example by making social comparisons or comparisons to previous self-states (12). In illuminating these judgements, researchers recognised elements of ‘response shift’ (Schwartz & Sprangers, 1999): adaptive changes, or recalibration of judgements about severity of limitation. The method also elucidated cognitions underlying risk judgements (6). In observing the information used to make judgements, think-aloud also allowed researchers to make inferences about personal models of susceptibility.

Four studies sought to explore what people think about when making judgements of HR-QoL at different points in time (16-17, 20-21). The method showed that people changed their standards of comparison over time. That is, they did not consistently anchor their QoL appraisals on the same characteristics over time. Over the course of chemotherapy treatment, respondents demonstrated changes in perspective (e.g. towards greater optimism) and self-presentation (e.g. presenting the self as coping). People also redefined what they perceived as important in their lives (20-21). This violated the assumption of consistency of cognitive processes underlying QoL appraisal (16-17). Participants often did not refer to their prior functioning when making judgements about changes in QoL, and many reports of prior functioning failed to correspond to those verbalised at pre-test (17). Likewise, some participants struggled to recall pre-test states (20). In this case, the results of think-aloud suggested a violation of the assumption of accurate recall of pre-test functioning, central to the measurement of QoL over time. Westerman et al. (2008) considered that in observing the self-assessment of QoL by means of think-aloud protocols, we have gained insight into the ‘black box’ of what actually happens in repeated QoL measurement.

In conclusion, spontaneous contributions made by respondents during think-aloud interviews allowed researchers insight into the way in which people interpret and respond to questionnaires; the models of health or susceptibility they hold; cognitions underlying decision or judgement-making; consistency in cognitive processes; and the frequency and nature of problems experienced with questionnaires. All authors considered that their findings uncovered important new knowledge with direct clinical and research implications.

Discussion

The literature suggests that think-aloud is a useful method to evaluate questionnaires in the health domain and all studies successfully used think-aloud to address their research aims. Specifically, think-aloud was used effectively to 1) assess the utility of questionnaires and to 2) understand cognitions involved in questionnaire completion. Justifications for using the approach were diverse, the main one being that it was used effectively in similar studies. Research findings were considered to evidence the efficacy of the method (4-5, 15, 22).

Think-aloud protocols were conducted effectively with participants of a range of ages, cognitive and physical capacity and level of education; with a variety of health-related questionnaires. Basic data can be obtained with a small sample (e.g. 6). Nevertheless, to ensure all issues are identified, sample size was best determined by the saturation of themes in the analysis. Where this is not feasible, Beatty and Willis (2007) recommend operating on a principle of diminishing returns. To statistically explore quantitative data, a larger sample size was required. Whilst the papers were selected on their use of concurrent think-aloud, almost half employed additional probe or interview techniques. Instruction protocols were most frequently based upon Green & Gilhooly, (1996), and many used a 'warm-up' procedure in line with best practice. Informative findings resulted from both quantitative and qualitative analyses.

In all papers, the think-aloud method met researchers' aims. The technique generated verbal data typically unobserved in the completion of questionnaire measures, allowing the evaluation of how well questions were meeting their objectives. This accords with the goal of cognitive interviewing paradigms (Beatty & Willis, 2007). The success of think-aloud in evaluating questionnaires may be attributed to a

number of advantages of the method. For instance, Willis (2005) suggests there is less interviewer-bias compared to other cognitive interviewing techniques, as the respondent is less affected by what he or she perceives to be important to the interviewer. Moreover, the interview may be closer to 'real life' survey completion, whereas interviewer probing can guide the direction and flow of the interview, creating artificiality (Conrad, Blair & Tracy, 2000). Think-aloud data is also collected during the response process which reduces reliance upon memory (Forsyth and Lessler, 1991; van der Veer, Hak, and Jansen, 2002). Other research would dispute the success of think-aloud in the current review. Nisbett and Wilson (1977) for example, question whether verbalisations are literal reflections of thought processes. Wilson, LaFleur and Anderson (1996) suggest that, to a degree, verbalisations probably reflect actual processes; however, they are more likely reconstructions. In contrast to Ericsson and Simon's (1980) assertion that thinking-aloud should not interfere with responding, Willis (1994) contends that the process increases the effort required to formulate a response. Moreover, Russo, Johnson, and Stephens (1989) found that thinking-aloud impacted the accuracy of various mental computations.

The results of the current review are in line with Drennan's (2003) review of cognitive interviewing in the process of pre-testing questionnaires. This paper specifically considered the use of cognitive interviewing (including think-aloud) in health care research, illustrating a number of studies which successfully used the technique to evaluate health-related questionnaires. Drennan (2003) provides a critique of the method, yet concludes that cognitive interviewing is of value in evaluating questionnaires, particularly in the pre-testing of questions which are sensitive; complex; in specific groups; or with respondents for whom questionnaire completion may present unique difficulties. The present review suggests moreover, that the evaluation of health-related questionnaires is indicated more widely, including with measures already in use; and that think-aloud interviewing is an effective means to this end. The findings of both reviews suggest that think-aloud could be applied to questionnaires outside of the health domain. Future research may also wish to extend the methodology beyond survey questionnaires, to more practical aspects of health care, such as in the exploration of processes underpinning medication or treatment adherence. For instance, to explore the extent to which

patients interpret medical advice or guidance in the manner intended by health care providers.

Strengths and limitations of the current review

The systematic nature of this review renders the procedure transparent and replicable, and the aims and method of the synthesis have been presented (Dixon-Woods, Agarwal, Jones, Young & Sutton, 2005). Literature was appraised for quality, ensuring that all papers included were of a sufficient standard to reliably inform the research questions, although no papers were excluded due to low quality. This demonstrates the existence of worthwhile research in this area. Analysis of the papers using a framework of predetermined features ensured that the research questions were answered. Different findings may have been obtained however, from a purely data-driven analysis. The search strategy (computerised databases only) precluded the inclusion of ‘grey’ literature. All studies included were published papers from peer-reviewed journals, again safeguarding the quality of findings, but opening the findings up to the possibility of bias arising from publication policies. For instance, it is conceivable that studies which used think-aloud unsuccessfully would not have been as readily published (Easterbrook, Berlin, Gopalan & Matthews, 1991).

Of note, the present review documents only the procedures and samples with which think-aloud has been used to date. Procedures or populations not described cannot be assumed to be ineffective or inappropriate. For instance, whilst papers excluded participants on the basis of being non-fluent in the native language of the questionnaire or interviewer, this does preclude these groups from successful participation in the future. For instance, Kudela et al. (2004) successfully implemented a multi-lingual think-aloud study of a tobacco use questionnaire across several languages, finding, at least in part, that some problems with survey questions were universal. This suggests then, not that think-aloud is unsuitable to evaluate questionnaires with foreign-language samples, rather, that to obtain a cross cultural and multi-lingual sample, may require greater resources in terms of interviewer language and interpretive skills.

Critically, the current review sought to elucidate the usefulness of the think-aloud method alone. This was not always possible however; in reporting results, authors

did not present the findings by method, resulting in the reporting of findings from an amalgamation of cognitive interviewing techniques. Notably, cognitive interviewing has been criticised for such opacity in the process of interviewing and analysis (Conrad & Blair, 1996). This is considered the principal weakness of the present review, and a legacy of the lack of clarity within the studies themselves. Future studies are advised to differentiate between the findings of each method used. Future research may also wish to compare the relative merits of cognitive interviewing methods used in alone or in combination; in particular, concurrent think-aloud protocols, follow-up probes and retrospective interviews. Absent from the current literature was consideration of the inherent acceptability of think-aloud interviewing in the evaluation of health-related questionnaires. This indicates a second direction for future research; for whilst the procedure may successfully address researchers' aims, it must present an acceptable burden to respondents to warrant its use.

Conclusions

Results indicate that think-aloud is a valuable tool for evaluating questionnaires in the health domain. The findings suggest that questionnaires currently in use are frequently not understood or completed in the manner intended by researchers. This may have far-reaching implications for our understanding of individual and population-wide health, delivery of interventions and health-care planning. Results suggest that further research into health-related questionnaires is warranted; moreover this can be effected successfully using think-aloud techniques.

References

- Al-Janabi, H., Keeley, T., Mitchell, P., & Coast, J. (2013). Can capabilities be self-reported? A think aloud study. *Social Science & Medicine*, 87, 116-122.
- Beatty, P. C., & Willis, G. B. (2007). Research synthesis: The practice of cognitive interviewing. *Public Opinion Quarterly*, 71, 1-25.
- Bode, C., & Jansen, H. (2013). Examining the Personal Experience of Aging Scale with the Three Step Test Interview. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences*, 9(3), 96-103.
- Brown, G. P., Hawkes, N. C., & Tata, P. (2009). Construct validity and vulnerability to anxiety: a cognitive interviewing study of the revised Anxiety Sensitivity Index. *Journal of Anxiety Disorders*, 23(7), 942-949.
- Critical Appraisal Skills Programme. (2001). 10 Questions to help you make sense of Qualitative Research. *Critical Appraisal Skills Programme: Making sense of evidence*. UK: Public Health Resource Unit
- Centre for Reviews and Dissemination. (2009). *Systematic Reviews: CRD's guidance for undertaking reviews in health care*. University of York, UK: CRD.
- Collins, D. (2003). Pretesting survey instruments: An overview of cognitive methods. *Quality of Life Research*, 12, 229-238.
- Conrad, F., & Blair, J. (1996). From impressions to data: Increasing the objectivity of cognitive interviews. *Proceedings of the section on survey research methods, vol. 1, pp. 1-9*. Alexandria, VA: American Statistical Association
- Conrad, F., Blair, J., & Tracy, E. (2000). *Verbal reports are data! A theoretical approach to cognitive interviews*. In Proceedings of the 1999 Federal Committee on Statistical Methodology Research Conference. Washington, DC: Office of Management and Budget.
- Converse, J. M., & Stanley, P. (1986). *Survey questions: Handcrafting the standardized survey questionnaire*. Newbury Park, CA: Sage Publications.
- Darker, C. D., & French, D. P. (2009). What sense do people make of a theory of planned behaviour questionnaire?: a think-aloud study. *Journal of Health Psychology*, 14(7), 861-871.
- DeMaio, T. J., & Landreth, A. (2004). Cognitive interviews: Do different methods produce different results? In S. Presser, J. M. Rothgeb, M. P. Couper, J. T.

- Lessler, E. Martin, J. Martin, & E. Singer (Eds.), *Methods for testing and evaluating survey questionnaires*. Hoboken, NJ: John Wiley and Sons.
- Dixon-Woods, M., Agarwal, S., Jones, D., Young, B., & Sutton, A. (2005). Synthesising qualitative and quantitative evidence: A review of possible methods. *Journal of Health Services Research & Policy*, 10, 45-53.
- Drennan, J. (2003). Cognitive interviewing: Verbal data in the design and pretesting of questionnaires. *J Adv Nurs*, 42:57-63
- Easterbrook, P. J., Berlin J. A., Gopalan R., & Matthews, D. R. (1991). Publication bias in clinical research. *Lancet*, 337, 867-72.
- Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. *Psychological Review*, 87, 215-250
- Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data, revised edition*. Cambridge, MA: MIT Press.
- Forsyth, B. H., & Lessler, J. T. (1991). Cognitive laboratory methods: A taxonomy. In P. P. Biemer, R. M. Groves, L. E. Lyberg, N. A. Mathiowetz, & S. Sudman (Eds.), *Measurement Error in Surveys*. Hoboken, NJ: John Wiley & Sons.
- Green, C., & Gilhooly, K. (1996). Protocol analysis: Practical implementation. In J. T. E. Richardson (Ed.), *Handbook of qualitative research methods for psychology and the social sciences* (pp. 55–74). Leicester, UK: BPS Books.
- French, D. P., Cooke, R., McLean, N., 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.
- French, D. P., & Hevey, D. (2008). What do people think about when answering questionnaires to assess unrealistic optimism about skin cancer? A think aloud study. *Psychology Health & Medicine*, 13(1), 63-74.
- Galasinski, D. (2008). Constructions of the self in interaction with the Beck Depression Inventory. *Health: an Interdisciplinary Journal for the Social Study of Health, Illness & Medicine*, 12(4), 515-533.
- Garcia, A. A. (2011). Cognitive interviews to test and refine questionnaires. *Public Health Nursing*, 28(5), 444-450.
- Hak, T., Veer van der, K., & Jansen, H. (2004). *The Three- Step Test-Interview (TSTI): An observational instrument for pre-testing self-completion*

- questionnaires*. Rotterdam, Netherlands: Erasmus Research Institute of Management. ERIM Report ERS-2004-029-ORG.
- Hak, T., van der Veer, C. J., & Jansen, H. (2008). The Three- Step Test-Interview (TSTI): An observation-based method for pretesting self-completion questionnaires. *Survey Research Methods*, 2, 143–150.
- Heesch, K. C., van Uffelen, J. G. Z., Hill, R. L., & Brown, W. J. (2010). What do IPAQ questions mean to older adults? Lessons from cognitive interviews. *The International Journal of Behavioral Nutrition and Physical Activity*, 7, 35.
- Horwood, J., Sutton, E., & Coast, J. (2013). Evaluating the face validity of the icecap-o capabilities measure: A "think aloud" study with hip and knee arthroplasty patients. *Applied Research in Quality of Life* Sep(Pagination), No Pagination Specified.
- Jansen, H., & Hak, T. (2005). The productivity of the Three- Step Test-Interview (TSTI) compared to an expert review of a self-administered questionnaire on alcohol consumption. *Journal of Official Statistics*, 21, 103–120.
- Jobe, J. B., & Mingay, D. J. (1989). Cognitive research improves questionnaires. *American Journal of Public Health*, 79(8), 1053–1055.
- Kaklamanou, D., Armitage, C. J., & Jones, C. R. (2013). A further look into compensatory health beliefs: a think aloud study. *British Journal of Health Psychology*, 18(1), 139-154.
- Kudela, M. S., Levin, K., Tseng, M., Hum, M., Lee, S., Wong, C., McNutt, S., & Lawrence, D. (2004). *Tobacco use cessation supplement to the current population survey Chinese, Korean, and Vietnamese translations: Results of cognitive testing*. Final report submitted to the National Cancer Institute, Rockville, MD.
- Lahaut, V. M. H., Jansen, H. A. M., & Garretsen, H. F. L. (2003). Comparison of two formats of the weekly recall and quantity-frequency alcohol measures. *Journal of Substance Use*, 8(3), 164-169.
- Mallinson, S. (2002). Listening to respondents: a qualitative assessment of the Short-Form 36 Health Status Questionnaire. *Social Science & Medicine*, 54(1), 11–21.
- McCorry, N. K., Scullion, L., McMurray, C. M., Houghton, R., & Dempster, M. (2013). Content validity of the illness perceptions questionnaire--revised

- among people with type 2 diabetes: a think-aloud study. *Psychology & Health*, 28(6), 675-685.
- McDowell, I. (2006). *Measuring health: A guide to rating scales and questionnaires* (3rd edit). New York, NY: Oxford University Press.
- Mulcahey, M. J., Chafetz, R. S., Santangelo, A. M., Costello, K., Merenda, L. A., Calhoun, C., . . . Betz, R. R. (2011). Cognitive testing of the spinal appearance questionnaire with typically developing youth and youth with idiopathic scoliosis. *Journal of Pediatric Orthopedics*, 31(6), 661-667.
- Nisbett, R. E., & Wilson T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, 84 231–59.
- Pool, J. J., Hiralal, S. R., Ostelo, R. W., van der Veer, K., & de Vet, H. C. (2010). Added value of qualitative studies in the development of health related patient reported outcomes such as the Pain Coping and Cognition List in patients with sub-acute neck pain. *Manual Therapy*, 15(1), 43-47.
- Popay, J., Roberts, H., Sowden, A., Petticrew, M., Arai, L., Britten, N., . . . Duffy, S. (2006). *Guidance on the conduct of narrative synthesis in systematic reviews: Final report*. Swindon, UK: ESRC Methods Programme.
- Russo, J. E., Eric J. J., & Stephens, D. L. (1989). The validity of verbal protocols. *Memory and Cognition* 17, 759–69.
- Schwartz, C. E., & Sprangers, M. A. G. (1999). Integrating response shift into health-related quality of life research: A theoretical model. *Social Science and Medicine*, 48, 1507–1515.
- Taminiau-Bloem, E. F., J, Visser, M. R. M., Tishelman, C., Schwartz, C. E., Koeneman, M. A., & Sprangers, M. A. G. (2011). Opening the black box of cancer patients' quality-of-life change assessments: A think-aloud study examining the cognitive processes underlying responses to transition items. *Psychology & Health*, 26(11), 1414-1428. doi: 10.1080/08870446.2011.596203
- Taminiau-Bloem, E. F., van Zuuren, F. J., Koeneman, M. A., Rapkin, B. D., Visser, M. R., Koning, C. C., & Sprangers, M. A. (2010). A 'short walk' is longer before radiotherapy than afterwards: a qualitative study questioning the baseline and follow-up design. *Health & Quality of Life Outcomes*, 8, 69.
- Tanur, J. M. (Ed.). (1992). *Questions about questions*. New York, NY: Russel Sage

- van der Veer, K., Hak, T., & Harrie, J. (2002). *The Three-Step Test Interview (TSTI): An observational instrument for pre-testing self-completion questionnaires*. Paper presented at the Questionnaire Development, Evaluation, and Testing Conference, Charleston, South Carolina.
- van Oort, L., Schroder, C., & French, D. (2011). What do people think about when they answer the Brief Illness Perception Questionnaire? A 'think-aloud' study. *British Journal of Health Psychology*, 16(2), 231-245.
- Waller, N. G. (1989). The effect of inapplicable item responses on the structure of behavioral checklist data: A cautionary note. *Multivariate Behavioral Research*, 24, 125-134.
- Watanabe, S., Nekolaichuk, C., Beaumont, C., & Mawani, A. (2009). The Edmonton symptom assessment system--what do patients think? *Supportive Care in Cancer*, 17(6), 675-683.
- Westerman, M. J., Hak, T., Sprangers, M. A., Groen, H. J., van der Wal, G., & The, A. M. (2008). Listen to their answers! Response behaviour in the measurement of physical and role functioning. *Quality of Life Research*, 17(4), 549-558.
- Westerman, M. J., The, A. M., Sprangers, M. A., Groen, H. J., van der Wal, G., & Hak, T. (2007). Small-cell lung cancer patients are just 'a little bit' tired: response shift and self-presentation in the measurement of fatigue. *Quality of Life Research*, 16(5), 853-861.
- Willis, G. B. (1994). *Cognitive interviewing and questionnaire design: A training manual (working paper series no. 7)*. Office of Research and Methodology at the National Center for Health Statistics, Centers for Disease Control and Prevention.
- Willis, G. B. (2005). *Cognitive interviewing: A tool for improving questionnaire design*. Thousand Oaks, Calif: Sage Publications.
- Willis, G., Royston, P., & Bercini, D. (1991). The use of verbal report methods in the development and testing of survey questionnaires. *Applied Cognitive Psychology* 5, 251-67.
- Wilson, T. D., LaFleur, S. J., & Anderson, D. E. (1996). The validity and consequences of verbal reports about attitudes. In N. Schwarz & S. Sudman (Eds.), *Answering questions: Methodology for determining cognitive processes in survey research*. San Francisco, CA: Jossey-Bass.

Wylde, V., Jeffery, A., Dieppe, P., & Gooberman-Hill, R. (2012). The assessment of persistent pain after joint replacement. *Osteoarthritis & Cartilage*, 20(2), 102-105.

Paper 2: Research Paper

Title

What sense do people with End Stage Kidney Disease make of the Emotion Regulation Questionnaire? A think-aloud study

The following paper has been prepared with submission to the 'British Journal of Health Psychology' in mind.

The guidelines for authors can be found in appendix A

Word Count: 6,820

Abstract: What sense do people with End Stage Kidney Disease make of the Emotion Regulation Questionnaire? A think-aloud study

Objectives. Researchers interested in the determinants of wellbeing in people with End Stage Renal Disease (ESRD) need reliable and valid instruments to assess emotional states and emotion regulation. However, there is no questionnaire designed specifically to assess emotion regulation in people with ESRD, a group who face many challenges including the need to undergo haemodialysis. The Emotion Regulation Questionnaire (ERQ) is a widely used instrument measuring cognitive reappraisal and emotional suppression, but we have no information on how it is understood in the ESRD population. The present study uses a “think-aloud” technique derived from cognitive interviewing methodology to examine the content validity of the ERQ with an ESRD population.

Design. A cross-sectional observational study.

Methods. 25 patients with ESRD who were undergoing hospital haemodialysis spoke their thoughts aloud as they completed the ERQ. Sessions were audiotaped and data subjected to content analysis.

Results. All questions were found to be problematic to some extent and reappraisal questions yielded the most problems. The most common problems were re-reading, mis-reading or floundering when answering questions (32 problems) and answering a different question or giving inconsistent reasoning (22 problems). A tendency to deny or minimise negative emotions and present a positive self-image was also noted.

Conclusions. Findings are indicative of problems with the ERQ when applied to an ESRD population. Caution is advised in using the ERQ with these patients. Replication is required with other populations to inform robust conclusions regarding the content validity of this measure.

Introduction

End Stage Renal Disease (ESRD) is the final stage (stage 5) of Chronic Kidney Disease (CKD). In this, the kidneys have ceased to function to such an extent that the patient is reliant on Renal Replacement Therapy (RRT) in the form of dialysis or kidney transplant. The prevalence of ESRD is estimated through the use of RRT, although this tends to be an underestimate as some patients will be managed conservatively. Worldwide, this has been identified as 316 people per million of the population (pmp), with an incidence of 73 per million. The prevalence is greater if we take high income countries where RRT is more accessible: 1283 pmp, with an incidence of 226 per million (Anand, Bitton & Gaziano, 2013); in England alone, the prevalence is 767 pmp. The aetiology of ESRD varies; however, 75% of cases are attributed to the effects of hypertension, glomerulonephritis and diabetes (Anand et al., 2013). In ESRD, mortality increases as kidney function decreases (Perazella & Khan, 2006), with death occurring most frequently from associated cardiovascular disease (e.g. Perazella & Khan, 2006; Tonelli, 2006).

ESRD can be managed with dialysis or kidney transplant. Dialysis mimics the functions of the healthy kidney by exchanging solutes and fluids and thereby purifying and detoxifying the blood artificially. The most common form of dialysis is haemodialysis in which the patient is attached to a dialysis machine for several hours; the alternative is peritoneal dialysis in which solutes and fluids are exchanged with dialysis solution internally through peritoneal membranes. Haemodialysis usually takes place in hospital 3 or 4 times per week, but can also be undertaken at home. Patients with ESRD receiving maintenance haemodialysis suffer from a multitude of physical symptoms, including fatigue, pain, muscle cramps, difficulty with sleep, and sexual dysfunction (Palmer, 2003; Rosas, 2001; Weisbord et al., 2005; Weisbord et al., 2007) and experience substantial impairments in quality of life (QOL) (e.g. Evans et al., 1985; Unruh, Weisbord & Kimmel, 2005). As noted above haemodialysis has to be carried out very regularly. For those patients on hospital haemodialysis this necessitates their attendance at a specialist Renal Unit on average three times weekly, interfering with employment and social participation. It

is unsurprising then, that a large epidemiological study found a near fourfold increase in depression in people with End Stage Renal Disease (ESRD) as compared to healthy individuals (Egede, 2007). Data on the prevalence and impact of anxiety in ESRD are scarcer; however, a review of 55 studies found 38% of people with ESRD experience substantial anxiety (Murtagh, Addington-Hall & Higginson, 2007). It is recognised that the affective concomitants of chronic illness have a significant impact upon health and wellbeing. For instance, in people with End Stage Renal Disease (ESRD), depression is associated with lower perceived health status and quality of life (Cukor, Coplan, Brown, Peterson & Kimmel, 2008); treatment non-adherence (Kimmel et.al., 1998); as well as increased mortality (e.g. Drayer et al., 2006). By implication, better understanding of the processes of emotional experience and management could enable the development and targeting of interventions with the potential to improve overall health status and life expectancy.

The process model of emotion regulation (Gross & Thompson, 2007) suggests that the way in which people regulate their emotions can influence affective experience. Two emotion regulation strategies have received particular attention. These are 'cognitive reappraisal' and 'expressive suppression' (Gross, 1998). 'Cognitive reappraisal' comprises attempts to think about a situation in a way which modifies its meaning and reduces emotional threat. This is an antecedent-focused strategy; it acts before an emotional response has been triggered (Gross, 2001; Gross & John, 2003). It may therefore be assumed to alter the entire temporal course of an emotional response. 'Expressive suppression' consists of attempts to minimise or inhibit the outward expression of emotion. By contrast, this is a response-focused strategy, intervening later in the course of the emotional response; once the emotional response is already underway (Gross, 2001; Gross & John, 2003). The two emotion regulation strategies are differentially adaptive; reappraisal is associated with more favourable consequences than suppression (Gross, 1998, 2001). Similar findings have been noted in studies of coping in chronic health conditions; for instance, reappraisal was associated with improved physical and psychological well-being in cardiac patients (Karademas, Tsalikou & Tallarou, 2011). Likewise, Gillanders, Wild, Deighan & Gillanders (2008) demonstrated the importance of emotion regulation in people with ESRD. Using the Emotion Regulation Questionnaire (ERQ; Gross & John, 2003), they found greater use of reappraisal to be positively

associated with the experience and expression of positive emotions, and negatively associated with the experience and expression of negative emotion. Reappraisal was also associated with lower levels of anxiety and a greater acceptance of their condition. By contrast, suppression was negatively associated with expression of positive emotion and positively associated with levels of depression, somatization, dissatisfaction with the support from others and dissatisfaction with the time spent managing their disease. These findings point to the need for further research into emotion regulation in an ESRD population. Evidence of the impact of emotion regulation strategies in ESRD, has the potential to inform interventions to promote the use of more adaptive strategies; ultimately, with the capacity to improve wellbeing, morbidity and mortality.

Use of the strategies of cognitive reappraisal and expressive suppression is commonly assessed using the Emotion Regulation Questionnaire (ERQ; Gross & John, 2003), a brief 10-question measure. Questions were derived from experimental research; 6 questions loading on the reappraisal factor and 4 on the suppression factor. Gross and John (2003) found good reliability for the scale (.79 for reappraisal; .73 for suppression) and a consistent two-factor structure across samples, ages, and cultures. They further evidenced the discriminant validity of the scale: for instance differentiating measures of reappraisal and suppression from cognitive ability, social desirability and personality. Reliability and validity results have been replicated with translated versions of the scale (e.g. Sala et al., 2012; Balzarotti, John & Gross, 2010) as well as with children and adolescents (Gullone & Taffe, 2012).

It is also vital to establish the content validity of questionnaires. That is, it is necessary to establish whether participants make sense of the questionnaire in the way intended by researchers; fundamentally, whether the questionnaire is measuring what it purports to assess. Establishing content validity is often undertaken using cognitive interviewing; of which concurrent 'think-aloud' interviewing was the original paradigm (Ericsson and Simon, 1980; Ericsson & Simon, 1993). In this method, participants are instructed to verbalise their thoughts as they read and complete a questionnaire. This is thought to allow insight into informants' underlying thinking and the contents of short term memory, in that whatever is consciously attended to is also verbalisable (Ericsson & Simon, 1993; Van den Haak, de Jong & Schellens, 2003). Importantly, participants are able to think-aloud

whilst simultaneously engaging in the target task: a review of over 40 think-aloud studies found no difference in task performance between those participants who verbalized their thoughts as compared those who completed the task in silence (Ericsson and Simon, 1993). This technique enables researchers to test the assumption, inherent in standardised surveys, that respondents are 1) universally able to understand the questions being asked of them; 2) that this understanding is the understanding intended by the researcher; and 3) that they are subsequently able, and willing, to answer such questions (Collins, 2003). A number of studies (Darker & French, 2009; French, Cooke, McLean, Williams & Sutton, 2007; Galansinski, 2008; van Oort, Schroder & French, 2011) have used think-aloud to examine questionnaires already in use in the health domain, for instance, the Theory of Planned Behaviour Questionnaire (TPB; see Ajzen, 1991, 2002); the Beck Depression Inventory (BDI; Beck, Steer & Brown, 1996); and the Brief Illness Perceptions Questionnaire (Brief IPQ; Broadbent, Petrie, Main, & Weinman, 2006) respectively. Although these scales are used frequently in the field, ‘thinking-aloud’ uncovered a number of problems when participants attempted to complete the questionnaires. The authors of the think-aloud studies subsequently made recommendations to improve the measures, or questioned the overall validity of the questionnaire for the intended purpose.

The ERQ is a frequently used measure and has previously been used to study emotion regulation in ESRD (Gillanders et al., 2008). To the best of our knowledge however, content validity has not been explicitly assessed for this measure. Specific to an ESRD population, interpretation of, and response to, the ERQ may be impacted by concomitants of their illness and treatment, such as pain, distress and fatigue. Given the existing burden of this group, it is vital that questionnaires are valid and that accurate inferences can be made from the data collected. The aim of the current study is to explore the processes of interpretation of, and responses to, the task of completing the ERQ with an ESRD population, using think-aloud interviewing. In exploring response processes, the present study will also examine verbalisations accompanying selection of the middle option on the response-scale (intended to denote a ‘neutral’ opinion). Previous research documents inconsistency in the reasoning given for selecting this option; ranging from outright refusal to answer, particularly in the case of requests for sensitive information, to an avoidance of

cognitive effort (Shoemaker, Eichholz & Skewes, 2002). Moreover, Darker and French (2009) found that participants were more likely to select the middle response-option for TPB questions experienced as more problematic. This was considered evidence of avoidance of effortful processing. In the case of the ERQ, selection of the middle response-option is interpreted by researchers as a neutral opinion. The present study aims to explore whether or not this assumption can be considered accurate.

Method

Participants

Participants were patients with ESRD undergoing hospital haemodialysis. Patients were recruited whilst on dialysis in the Specialist Renal Dialysis Units of a large hospital in the UK. Inclusion criteria required participants to be aged 18-80 and to be fluent in reading and speaking the English language. In total 35 patients were approached; 8 declined to participate and 27 were recruited to the study. One was subsequently withdrawn due to difficulties with literacy; a second was withdrawn as thoughts were not verbalised. Twenty-five patients, 20 males, with a mean age of 56.64 years (s.d.= 17.53, range – 20-77) therefore completed the study.

Measures

The think-aloud procedure was conducted with The Emotion Regulation Questionnaire, a 10-question scale to measure habitual use of cognitive reappraisal and emotional suppression. Questions were presented as statements, for instance: "I control my emotions *by changing the way I think* about the situation I'm in" (reappraisal); and "I control my emotions *by not expressing them*" (suppression). In addition to these general-emotion questions, both scales included at least one question which referred to regulating negative emotions (e.g. *sadness* and *anger*) and one question about regulating positive emotion (e.g. *joy* and *amusement*). Responses were made on a 7-point Likert scale from 'strongly disagree' to 'strongly agree'. Additional information was gathered to describe the demographic and clinical characteristics of the sample. (See table 2).

Design and Procedure

The study employed a cross-sectional, cognitive interviewing design, using a think-aloud procedure. All participants agreed to be audio-recorded and informed consent

was obtained. The study was granted ethical approval from the NRES Committee Northwest Liverpool East.

Interviews were conducted on the Renal Unit whilst participants were undergoing haemodialysis. Participants were interviewed individually, and in private, by the first author. To detail the think-aloud procedure, participants were read the following instructions, adapted from Green and Gilhooly (1996) and French et al. (2007):

We are interested in how people regulate their emotions. There is a questionnaire which measures the way that people regulate emotions. We want to check that people understand the questions in the way that we meant them. To do this, I am going to ask you to think-aloud as you complete the questionnaire. What I mean by 'think-aloud' is that I want you to tell me everything you are thinking as you read each question and decide how to answer it. I would like you to talk aloud constantly. I don't want you to plan out what you say or try to explain to me what you are saying. Just act as if you are alone in the room speaking to yourself. If you are silent for any long period of time, I will ask you to talk. Please try to speak as clearly as possible, as I shall be recording you as you speak. Do you understand what I want you to do?'

Participants were not interrupted during thinking-aloud, unless they fell silent for 10 seconds, in which case they were asked to 'keep talking'. In accordance with recommendations for best practice (Green and Gilhooly, 1996) a 'warm-up' task was completed prior to the ERQ. Five questions from the Illness Perceptions Questionnaire-Revised (IPQ-R; Moss-Morris et al., 2002) were used for this purpose. Any queries were dealt with at this point and misunderstandings corrected.

Analysis

Each interview was transcribed verbatim. Transcripts were segmented into verbalisations relating to each of the 10 questions of the questionnaire. Segments were examined alongside the completed questionnaire and coded according to whether or not any difficulties were articulated with the questions. Initially coding categories were based upon those of French et al. (2007):

- 1) No significant problems identified.

- 2) Participants re-read question, mis-read question, or significantly flounder in answering it. Participants were aware that they experienced problems in understanding the question.
- 3) Questioned sensibleness of question. Respondent identified problem with how question was worded.
- 4) Answered different question from the one that was asked, or gave reasoning inconsistent with the answer given. Participants had problems in comprehending/answering question but were not themselves necessarily aware of this.

in addition to one code from McCorry et al. (2013):

- 5) Participant expressed that question or response-options were not applicable to their circumstances.

Inspection of the data indicated the need for a further category to reflect difficulties with the scale only:

- 6) Struggle with comprehension or use of the scale.

All transcripts were coded by the first author. Specifically, for each participant in turn, each question was coded (0,1) as to whether or not each type of problem was present. Therefore, for each question, the number of problems for each participant was noted. For each question, the total number of problems experienced across all participants was calculated. The number of participants having any problems (as opposed to no problems) with each question was also calculated, and the number of problems experienced by each participant was calculated. Transcripts were independently coded by a second researcher (using the coding scheme described above). This found 36 disagreements out of 250 responses coded, giving an agreement rate of 85.6%. In this instance, there were too few participants and too many different response combinations to calculate a Cohen's kappa.

A separate content analysis was conducted by the first author to identify 1) participants' views about the questionnaire (from spontaneous verbalisations) and 2) rationales given for selecting the middle response-option (from verbalisations accompanying a 'neutral' scale response). Views regarding the questionnaire were

coded as ‘problematic’ or ‘unproblematic’. For the middle response-option, four categories of rationale were observed in the data (see table 6).

Finally, a thematic analysis was conducted to explore verbalisations related to emotional experience more generally. Thematic analysis was conducted upon limited features of the data set, that is, upon spontaneous comments related to participants’ experiences of emotion. 19 out of 25 participants spontaneously made comments related to emotion. Thematic analysis was performed by the first author, progressing through the six phases defined by Braun and Clarke (2006). In phase one, the first author became familiarised with the data through the process of interviewing, reading and re-reading transcripts and performing the above content analyses. Spontaneous comments related to participants’ experiences of emotion were noted and considered independently from the rest of the data which consisted of comments related to the questionnaire (this was content analysed as described above). In phase two, initial codes were developed to reflect interesting aspects of all of the most basic segments of raw data. Example codes are: ‘stay positive for others’; ‘put on a brave face’; and ‘denial of negative emotion’. In phase three, themes were identified from the codes and, in an iterative process, themes were considered and reconsidered in relation to emerging themes and all relevant codes collated into the themes. These themes were refined in phase four. In this, all coded extracts were reviewed in relation to each theme; on the basis of coherence within the theme, the themes were then reviewed in relation to the entire data set. This determined whether the themes appropriately represented the data and searched to code additional data within the themes. In phase 5, the themes were defined in terms of what aspects of data they represent, and a name selected to capture the ‘essence’ of the theme. In Phase 6, the results of the analysis were reported in ‘Results’, ‘*Spontaneous emotion-related comments*’.

Results

Sample characteristics

Table 2

Participant Demographics, Relationship Status, Education and Employment Status and Dialysis-Related Variables

Demographic Characteristics					
Male	Female				
20	5				
Age					
18-30	30-40	40-50	50-60	60-70	70-80
1	6	0	5	4	9
Ethnicity					
White, British	Irish	Indian	Caribbean	Other Asian	
20	1	1	1	1	
Relationship Status					
Single	Married/ cohabiting	Separated/ divorced	Widowed		
10	12	1	2		
Education					
No qualifications	GCSE's, CSE's or O-levels	A levels/ BTEC	Trade/ Apprenticeship	University degree	Other
5	7	4	6	4	2
Employment Status					
Full time	Part time	Home duties	Unemployed	Retired	
5	1	0	6	13	
Reported energy level					
‘Tired’	‘A bit tired’	‘Not tired’			
5	4	15			
Number of days since last had haemodialysis					
1	2	3			

2	22	1	
Length of time on dialysis at interview (hours)			
<0.5	0.51-1.5	1.51-2.5	2.51+
8	4	6	7

Overall, the 25 participants experienced a total of 93 problems with the 10 ERQ questions. The median number of problems experienced by participants was 3.5 (mean = 3.7, range = 0-9, s.d. = 2.34) with only one participant having no problems.

Distribution of problems across the questions

Table 3 demonstrates the number of participants who were coded as having any type of problem for each question. It should be noted that participants may be coded as having more than one problem with each question. None of the questions were completely unproblematic and some questions were more problematic than others. The most problematic was question 3 ‘When I want to feel less *negative* emotion (such as sadness or anger), I *change what I’m thinking about*’. 14/25 (56 %) of participants were coded as having problems with this question. Question 1 ‘When I want to feel more *positive* emotion (such as joy or amusement), I *change what I’m thinking about*’ was the second most problematic question; evidenced by difficulties in 13/25 (52%) of participants. The least problematic was question 2 ‘I keep my emotions to myself’, which elicited difficulties in only 2/25 (8%) of participants. The remaining questions were problematic in 5-10/25 (20-40%) of participants. More problems were experienced with the cognitive reappraisal questions (mean number of problems encountered per question 12) than with the suppression questions (mean number of problems encountered per question, 5.2).

Table 3

Number of Participants Coded as Having Any Type of Problem for Each Question

Number participants coded as having any type of problem for each question (Q1-Q10)									
Q1 (R)	Q2 (S)	Q3 (R)	Q4 (S)	Q5 (R)	Q6 (S)	Q7 (R)	Q8 (R)	Q9 (S)	Q10 (R)
13	2	14	6	9	5	9	7	7	10

Note. R = Reappraisal question; S = Suppression question

Nature of the problems identified

The most common problems coded were 1) re-reading, misreading, or floundering in answering (comprising 32 of the 93 problems) and 2) answering a different question from that asked, or giving inconsistent reasoning (22/93 problems). The least

frequent issue was ‘questioning the sensibleness of the question’ (6/93 problems). Verbalisations are quoted to exemplify problems (table 4).

Table 4. *Frequency with which each Type of Problem was Coded and Extracts to Illustrate*

Problem	Frequency	Example Quotes
1) Re-read, mis-read, or significantly flounder in answering question	32	“When I want to feel more positive emotion I change the way, what I am thinking about erm change the way I’m thinking about the situation .. erm think I’ll have to put a neutral on that, ‘cos I’ve no idea what that means” “question 9 when I am feeling negative emotions I make sure not to express them .. er, (laughs) .. when I am feeling negative emotions I make sure not to express them .. erm .. I’m not too sure about that one .. when I’m feeling negative emotions I make sure not to express them erm sorry, I’m gonna come back to that question”
2) Questioned sensibleness of question	6	“Well that’s a bit vague actually that one I change what I’m thinking about ..” “that’s .. kind of the same question ..”
3) Answered different question from the one that was asked, or gave inconsistent reasoning	22	“well I do, I’ll ring somebody up or nip round me friends or something, go and get me hair done. I change what I’m thinking about, so that’s, erm, 7 again.” “if that means that I .. think about the situation to help deal with it, then I strongly agree”

4) Question deemed not applicable to their circumstances	16	“I don’t, because usually I’ve never really felt negative ..” “I don’t feel negative, I don’t believe in negative .. might as well get in a box if you’re gonna be negative”
5) Problems with the response-scale	17	“oh, I can’t say I change so where do I go with that one? .. neutral? .. if I don’t change?” “sometimes”

The nature of the problems varied between the questions. Notably, participants experienced difficulties with the response-scale twice as often for question 1 as compared the other 9 questions; this and 2) re-reading, misreading or floundering in answering, constituted the main difficulties with this question. Also noteworthy is that the reappraisal questions were coded as ‘not applicable to me’ 15 times as compared only once for the suppression questions. Table 5 displays the frequency and nature of problems for each question.

Table 5

Frequency of each Type of Problem Coded for Each Question on the Emotion Regulation Questionnaire

	Number of times each problem type was coded as present for each question (Q1-Q10)										Total
	Q1 (R)	Q2 (S)	Q3 (R)	Q4 (S)	Q5 (R)	Q6 (S)	Q7 (R)	Q8 (R)	Q9 (S)	Q10 (R)	
Re-read, mis-read or flounder in answering	5	1	3	2	3	2	5	3	3	5	32

Question sensibleness of question	1	0	2	0	0	0	1	1	0	1	6
Answered different question or gave inconsistent reasoning	3	0	4	3	3	1	2	1	3	2	22
Not applicable to me	1	0	5	0	3	0	1	0	1	5	16
Struggle with scale	6	1	2	1	0	2	0	3	1	1	17
Total number of problems	16	2	16	6	9	5	9	8	8	14	93

Note. R = reappraisal question; S = Suppression question

Although on average cognitive reappraisal questions were over twice as problematic, the types of problems reported appeared to be similarly distributed across the reappraisal and suppression questions (see table 6). The exception was problem 5) ‘not applicable to me’, which was applied 15 times to the 6 reappraisal questions and only once to the 4 suppression questions. Participants exhibited this difficulty a mean of 2.5 times for reappraisal questions compared to 0.25 times for suppression questions.

Table 6

Mean Frequency with which each Type of Problem was Coded for Cognitive Reappraisal Questions (6) and Emotional Suppression Questions (4)

Mean frequency with which each type of problem was coded as present

Problem	Reappraisal (6 questions)	Suppression (4 questions)
Re-read, mis-read or flounder in answering	4	2
Question sensibleness of question	1	0
Answered different question or gave inconsistent reasoning	2.5	1.75
Not applicable to me	2.5	0.25
Struggle with scale	2	1.25
Total number of Problems	12	5.25

Participants' Opinion of the ERQ.

10 views were expressed to the effect that the questionnaire was perceived as “okay” or “easy to use”; these were coded as ‘participants experienced the questionnaires as unproblematic’. To a greater extent however, comments were coded as ‘participants found the questionnaire problematic’ (19 verbalisations). The predominant issue (articulated 7 times) was with the similarity or repetitiveness of the questions. Other types of problems were diffuse, but included finding the questionnaire “confusing” (3 participants), “irrelevant” (1 participant), or perceiving that “the questions seem to contradict each other” (3 participants).

Rationale for selecting the middle response-option

Participants varied in the rationales they gave for selecting the middle response-option. There were judged to be 4 categories to the rationales provided (table 4). Participants most frequently (36% of verbalisations) indicated that in selecting the middle option, they did not understand the question or scale; only 21% of verbalisations indicated that selection of the middle response was indicative of a ‘neutral’ opinion (code 3, table 7).

Table 7

Participants' Rationales for Selection of the Middle Response-Option: Frequency with which each Rationale was Coded and Extracts to Illustrate

Code	Count	Example
------	-------	---------

1. 'Does not apply to me'	6	"I don't know I've no anger .. I don't have no anger anger, anger .. I'll put .. 4"
2. 'Depends on the situation/degree of emotion'	6	"depends on the conditions, situation, so I'm gonna put 4 on that one"
3. 'No strong feelings with regard the question or agree 'a little bit''	6	"I don't have any particular feelings on that I'd say I feel neutral on that"
4. 'Does not understand the question/scale'	10	"I'll have to go 4 on that neutral, I don't know what that's, don't even know what that means to be honest .."

Spontaneous emotion-related comments

Thematic analysis identified two main themes within verbalisations: 'denial of negative emotion' and 'presenting a positive self-image'. With regard the former, many raw data segments were coded as 'denial of negative emotion' in phase 2 of thematic analysis. This code went on to form a main theme. This theme was retained after review in phases 4 and 5 of analysis. This demonstrates internal homogeneity of the theme in line with Patton's (1990) criteria for judging categories. To illustrate this theme, many people asserted that they did not experience negative emotions: "I don't get negative emotions"; "I don't get stressed". Some referred also referred to specific negative emotions: "I've no anger .. I don't have no anger anger, anger..."; or to 'stress': "I don't believe in stress"; "I don't get stressed...". This theme appears to represent participants' sense of their emotional life as predominantly positive; or at least, it represents the sense of themselves which they were willing to share with the interviewer. This is important in terms of the wider research aim: to explore the processes of interpretation of, and responses to, the ERQ in an ESRD population. Denial of negative emotion in this sample, may render the questions on the ERQ related to negative emotions less relevant or applicable. Whilst this theme predominated, there were also a few comments to the contrary. These suggested more of an acceptance of negative emotion: "I've always lived with sadness and anger and it's just part...".

The latter theme 'presenting a positive self-image', was characterised by presenting the self as positive, stoical, a 'coper'. This was an over-arching theme comprised of

codes identified at phase 2, such as ‘stay positive; ‘don’t express negative emotions’; ‘look on the ‘bright side’; and ‘put on a ‘brave face’’. This theme was initially named ‘staying positive’, and was refined in phases 4 and 5 to ‘presenting a positive self-image’ to more accurately reflect the ‘essence’ of the data. That is, to reflect the importance of staying positive in relation to another. Many people articulated statements to this effect: “I cope by keeping it in – be happy to people”; “I like to stay positive”; “I am a positive person though, I am, everybody says so...”; “I generally try and look on the bright side of life” ; “I’m always very upbeat” ; “I’ve always got to try and put a happy slant on things”. This theme captures the importance of presenting the self in a positive light. Again, this likely impacts the processes of interpretation of, and responses to, the ERQ in the presence of an interviewer. Again, to a lesser extent, there were exceptions to this theme. A few individuals indicated that they did express emotions: “I do bloody well express ‘em (laughs)”; and some admitted expressing emotion under exceptional circumstances: “I think sometimes it’s good to ...let out a bit of sadness or even a bit of anger sometimes”.

Discussion

This is the first study in which people have thought aloud whilst completing the ERQ. Participants were adults with ESRD on hospital haemodialysis. Participants did not have difficulties with two thirds of the questions; however, a variety of problems were found with the questionnaire as a whole. These findings are similar to those found by Darker & French (2009). All participants, bar one, experienced at least one type of problem with at least one question; the average number of problems experienced by participants being 3.7. The most problematic questions were those pertaining to cognitive reappraisal. Question 2 (a suppression question) was the least problematic. No question was completely unproblematic and the nature of problems varied between the questions. More often than not, participants expressed the opinion that the questionnaire was problematic. A further finding was that selection of the middle response-option was rationalised as a ‘neutral’ opinion in only one fifth of verbal accounts, suggesting that interpretations based on this assumption may

often be incorrect. Of note the majority of participants sampled presented an image of themselves as ‘positive’, a ‘coper’ and denied the experience of negative emotions. These characteristics may have impacted the way in which subjects interpreted and responded to the ERQ.

Distribution and nature of problems with the ERQ

It is notable that the most problematic questions were all questions pertaining to the use of cognitive reappraisal. This suggests that reappraisal questions are either: 1) systematically expressed in a way which is harder to interpret and respond to; or that 2) the concept of cognitive reappraisal is implicitly more challenging. The former may be ameliorated with adaptations to the wording of questions; the latter, however, would have more serious implications for the validity of the scale as a whole. French et al. (2007) make a similar distinction when appraising the tractability of problems arising with the TPB Questionnaire. In support of hypothesis (1) it should be noted that the average number of words per question is 18 for reappraisal questions compared to 10 for suppression. Furthermore the 4 most problematic questions (all reappraisal questions) contain double negative or double positive statements (e.g. question 3: ‘When I want to feel *less negative* emotion (such as sadness or anger), I change what I’m thinking about). McCorry, Scullion, McMurray, Houghton & Dempster (2013) consider that the inclusion of such questions can add a degree of cognitive complexity to the task of completing a questionnaire. On balance, they judged that negatively worded questions could result in more problems than they solve and should be avoided in the construction of questionnaires (Roszkowski & Soven, 2010). Interestingly, the problems coded for question 1 were predominantly difficulties with the scale and re-reading. These may be attributable to presentation bias as people become familiarised with the measure, rather than problems with the question per se. This said, participants considered that reappraisal questions were not applicable to their circumstances 15 times as frequently as suppression questions, suggesting more difficulties with the concept of cognitive reappraisal overall. Future research is required to explore the impact of adapting the wording of reappraisal questions.

Most participants voiced the opinion that the ERQ was problematic. Whilst a number of comments referred to the repetitiveness of the questions, other complaints were diffuse. This suggests that on an individual-basis people may struggle with certain

aspects of the scale; it does not however, indicate systematic problems with the measure. Consistent with previous literature (Willis, Royston & Bercini, 1991) it was striking that participants completed all questions, regardless of whether or not questions were perceived as applicable to the respondent, or even understood. This is in line with evidence from experimental research showing that survey participants try to answer questions no matter how difficult they find it (Tanur, 1992).

Of note, the majority of participants in the present study denied experiencing negative emotions and sought to present a positive self-image. This may explain why a number of participants thought that questions were not applicable to their circumstances (for example, question 3 ‘When I want to feel less *negative* emotion (such as sadness or anger), I *change what I’m thinking about*’). This phenomenon may perhaps be attributable to the sample, consisting to a large extent, of older male participants. Potentially these participants found it socially unacceptable to express negative emotions in the presence of a younger, female interviewer. Masculinity is generally associated with presenting the self as tough and unemotional; avoiding displays of weakness or dependency (Knobloch & Metts, 2013). Previous research found emotional suppression to be adaptive for men; for instance, less expression of negative emotion has been associated with higher social status for men but not for women (Anderson, John, Keltner, & Kring, 2001). The views expressed in the current study could also be a consequence of managing a serious chronic health condition such as ESRD. For instance, having a stoic attitude is considered one way in which older people cope with the effects of chronic pain (Cairncross, Magee & Askham, 2007; Helme & Gibson, 1999). Nonetheless, these are precisely the kinds of questions a measure such as the ERQ may be hoped to elucidate. It would be interesting to examine whether there would be a reduction in questions coded as ‘not applicable to me’ with the introduction of an initial screening questionnaire. Administered prior to the ERQ, this might first clarify whether participants experience a range of positive or negative emotions. This may further act to familiarise participants with the concepts of emotion raised in the ERQ.

Interpretation of the middle response-option

The current study indicated that the selection of ‘neutral’ on the ERQ response-scale cannot be assumed to indicate a ‘neutral’ opinion of the question. For some participants it appeared that selecting the middle option was a way of dealing with

questions perceived as not applicable to them; for others, ‘neutral’ was selected when none of the other response-options were deemed appropriate. The greatest proportion of participants selected neutral when they did not understand the question or scale. Consistent with previous research (Shoemaker, Eichholz & Skewes, 2002; Darker and French, 2009), the latter seems indicative of an avoidance of effortful processing.

Strengths and limitations of the current study

The study is strong if the focus of interest is upon the measurement of emotion regulation in ESRD patients using the ERQ. People with ESRD must constantly manage the physical and psychosocial burden of their disease and treatment. It is thus imperative that further burden placed upon them, through research or outcome measurement, is valid and can accurately inform knowledge and health interventions. The ERQ has been used across a range of medical conditions including ESRD; however, the validity of the questionnaire has not been established with these groups. This study addresses this issue and is strong in that it suggests caution in the use of the ERQ with this group. Potentially this means that people with ESRD will not be burdened with a questionnaire which may have limited validity in this population.

The ESRD sample may however, reduce generalisability to other populations. Whilst the present findings highlight issues with the ERQ, caution is advised when generalising findings to other populations. In addition, the sample was predominantly male; as discussed, this may have implications for the way in which the participants viewed and presented their emotions; and the way in which they interpreted and responded to the questionnaire. The sample size was relatively small; however think-aloud methodology has been successfully used to assess questionnaires with as few as 6 participants (French et al., 2007; van Oort, Schroder & French, 2011). Further, interviews were transcribed verbatim, resulting in rich data.

A weakness of the present study is that the analysis does not differentiate difficulties which result in response problems, from those which do not. Future research may wish to make this distinction in order to elucidate the seriousness of the problems identified. Potentially, problems which do not result in clear response problems (such as re-reading) could be coded as ‘response error-free’; whereas, problems such

as answering a different question from the one asked, or giving inconsistent reasoning, would be coded as a 'response error'. This is a crude measure however; response errors would only be detected to the extent that they are clearly articulated. Galasinski (2008), considers that a code of 'no significant problem' can only be understood as the absence of verbalised difficulties; not the presence of an accurate and appropriate response. The implication for the current study is that problems may actually be underrepresented.

Think-aloud methodology was successfully used in the present study to uncover processes of interpretation and response to the ERQ. With the exception of one person who was withdrawn from the study, participants managed to think-aloud whilst completing the ERQ. In accordance with previous research, think-aloud techniques identified the nature and frequency of problems with the questionnaire (e.g. French et al., 2007; Horwood, Sutton & Coast, 2013) and in using the response-scale (Darker & French, 2009). A limitation of the technique, however, is the reliance upon participants verbalizing their thoughts; problems which are not articulated are therefore missed. A further criticism relates to the artificial experience of thinking aloud in the presence of an interviewer. This said, Willis (2005) perceives there to be less interviewer-bias from think-aloud methods as compared to other cognitive interviewing techniques as the respondent is less affected by what he or she perceives to be important to the interviewer. A number of other studies also consider think-aloud to introduce less distraction or social desirability bias (e.g. Darker & French, 2009; Horwood et al., 2013; Mallinson, 2002) and to provide the most naturalistic method of enquiry. Nonetheless, the use of additional cognitive interviewing techniques, such as verbal probes, may have facilitated the clarification of verbalisations in the current study (Willis, 1991).

The present study would benefit from replication with a larger and more representative sample, as well as with other clinical and non-clinical populations, to explore whether similar patterns of difficulties emerge. Currently the findings are indicative of potentially problematic aspects of the ERQ; however, results would need to be replicated in more representative samples before robust conclusions could be made about the validity of the measure.

References

- Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, 50, 179–211.
- Ajzen, I. (2002, September). Constructing a TPB questionnaire: Conceptual and methodological considerations. Retrieved from http://chuang.epage.au.edu.tw/ezfiles/168/1168/attach/20/pta_41176_7688352_57138.pdf
- Anderson, C., John, O. P., Keltner, D., & Kring, A. M. (2001). Who attains social status? Effects of personality and physical attractiveness in social groups. *Journal of Personality and Social Psychology*, 81, 116–132.
- Anand, S., Bitton, A. & Gaziano, T. (2013). The Gap between estimated incidence of renal disease and use of therapy. . PLoS ONE 8(8): e72860. doi:10.1371/journal.pone.0072860
- Balzarotti, S., John, O. P., & Gross, J. J. (2010). An Italian adaptation of the Emotion Regulation Questionnaire. *European Journal of Psychological Assessment*, 26, 61-67.
- Beck, A.T., Steer, R.A., & Brown, G.K. (1996). *Beck Depression Inventory II.Manual*. San Antonio, TX: The Psychological Corporation.
- Braun, V., & Clarke, V. (2006) Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77-101.
- Broadbent, E., Petrie, K. J., Main, J., & Weinman, J. (2006). The Brief Illness Perception Questionnaire. *Journal of Psychosomatic Research*, 60, 631–637 doi:10.1016/j.jpsychores.2005.10.020
- Cairncross L., Magee H., & Askham L. (2007). A hidden problem: Pain in older people: A qualitative study. Report. Oxford, UK: Picker Institute Europe.
- Collins, D. (2003). Pretesting survey instruments: An overview of cognitive methods. *Quality of Life Research*, 12, 229–238
- Conrad, F., Blair, J., & Tracy, E. (1999). *Verbal reports are data! A theoretical approach to cognitive interviews*. Federal Committee on Statistical Methodology Research Conference, Arlington, USA.
- Cukor, D., Coplan, J., Brown, C., Peterson, R. A., & Kimmel, P. L. (2008). Course of depression and anxiety diagnosis in patients treated with hemodialysis: a 16-month follow-up. *Clinical Journal American Society Nephrology*, 3, 6, 1752-1758.

- Darker, C. D., & French, D. P. (2009). What sense do people make of a theory of planned behaviour questionnaire?: A think-aloud study. *Journal of Health Psychology, 14*(7), 861-871.
- Drayer, R. A., Piraino, B., Reynolds, C. F., 3rd, Houck, P. R., Mazumdar, S., Bernardini, J., et al. (2006). Characteristics of depression in hemodialysis patients: Symptoms, quality of life and mortality risk. *Gen Hosp Psychiatry, 28*, 4, 306-312.
- Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. *Psychological Review, 87*, 215-250
- Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data, revised edition*. Cambridge, MA: MIT Press.
- French, D. P., Cooke, R., McLean, N., 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.
- Galasinski, D. (2008). Constructions of the self in interaction with the Beck Depression Inventory. *Health: an Interdisciplinary Journal for the Social Study of Health, Illness & Medicine, 12*(4), 515-533.
- Gillanders, S., Wild, M., Deighan, C., & Gillanders, D. (2008). Emotion regulation, affect, psychosocial functioning, and well-being in hemodialysis patients. *Am J Kidney Dis. 51*, 4, 651-62.
- Green, C., & Gilhooly, K. (1996). Protocol analysis: Practical implementation. In J. T. E. Richardson (Ed.), *Handbook of qualitative research methods for psychology and the social sciences* (pp. 55–74). Leicester, UK: BPS Books.
- Gross, J.J. (1998). Antecedent- and response-focused emotion regulation: Divergent consequences for experience, expression, and physiology. *Journal of Personality and Social Psychology, 74*, 224–237.
- Gross, J.J. (2001). Emotion regulation in adulthood: Timing is everything. *Current Directions In Psychological Science, 10*, 214–219.
- Gross, J.J., & John, O.P. (2003). Individual differences in two emotion regulation processes: Implications for affect, relationships, and well-being. *Journal of Personality and Social Psychology, 85*, 348–362.
- Gross, J.J., & Thompson, R.A. (2007). Emotion regulation: Conceptual foundations. In J.J.

- Gross (Ed.), *Handbook of emotion regulation*. New York, NY: Guilford Press.
- Gullone & Taffe (2012). The Emotion Regulation Questionnaire for Children and Adolescents (ERQ-CA): A psychometric evaluation. *Psychological Assessment* 24, 2, 409–417. doi: 10.1037/a0025777
- Helme R.D., & Gibson S.J. (1999). Pain in older people. In: I. K. Crombie (Ed.), *Epidemiology of Pain* (pp. 103–112). Seattle, Wash: IASP Press.
- Horwood, J., Sutton, E., & Coast, J. (2013, September). Evaluating the face validity of the ICECAP-O capabilities measure: A "think aloud" study with hip and knee arthroplasty patients. *Applied Research in Quality of Life*. Retrieved from <http://link.springer.com/article/10.1007%2Fs11482-013-9264-4#page-1>
- Karademas, E. C., Tsalikou, C., & Tallarou, M. C. (2011). The impact of emotion regulation and illness-focused coping strategies on the relation of illness-related negative emotions to subjective health. *J Health Psychol* 16, 3, 510-9.
- Kimmel, P. L., Peterson, R. A., Weihs, K. L., Simmens, S. J., Alleyne, S., Cruz, I., & Veis, J. H. (1998). Psychosocial factors, behavioural compliance and survival in urban haemodialysis patients. *Kidney Int*, 54, 1, 245-254.
- Knobloch, L.K., & Metts, S. (2013). Emotion in relationships. In J.A. Simpson and L. Campbell (Eds.), *The Oxford handbook of close relationships* (pp. 514-534). Oxford, England: Oxford University Press.
- Mallinson, S. (2002). Listening to respondents: a qualitative assessment of the Short-Form 36 Health Status Questionnaire. *Social Science & Medicine*, 54(1), 11–21.
- McCorry, N. K., Scullion, L., McMurray, C. M., Houghton, R., & Dempster, M. (2013). Content validity of the illness perceptions questionnaire--revised among people with type 2 diabetes: a think-aloud study. *Psychology & Health*, 28(6), 675-685.
- McRae, K., Jacobs, S.E., Ray, R.D., John, O.P., & Gross, J.J. (2012). Individual differences in reappraisal ability: Links to reappraisal frequency, well-being, and cognitive control. *Journal of Research in Personality*, 7, 253-262.
- Moss-Morris, R., Weinman, J., Petrie, K. J., Horne, R., Cameron, L. D., & Buick, D. (2002). The revised Illness Perception Questionnaire (IPQ-R). *Psychology & Health* 17, 1, 1–16.

- Perazella, M., A, Khan, S. (2006). Increased mortality in chronic kidney disease: a call to action. *The American Journal of the Medical Sciences*, 331, 3, 150–3. doi:10.1097/00000441-200603000-00007. PMID 16538076.
- Roszkowski, M.J., & Soven, M. (2010). Shifting gears: Consequences of including two negatively worded items in the middle of a positively worded questionnaire. *Assessment & Evaluation in Higher Education*, 35, 117–134. doi: 10.1080/02602930802618344
- Sala, M.N., Molina, P., Abler, B., Kessler, H., Vanbrabant, L., & van de Schoot, R. (2012). Measurement invariance of the Emotion Regulation Questionnaire (ERQ). A cross- national validity study. *European Journal of Developmental Psychology*. 9:6, 751-757. doi:10.1080/17405629.2012.690604
- Shoemaker, P. J., Eichholtz, M., & Skewes, E, A. (2002). Item nonresponse: Distinguishing between don't know and refuse. *International Journal of Public Opinion Research*, 14, 2, 193-201
- Tanur, J. M. (Ed.). (1992). *Questions about questions*. New York: Russel Sage
- Tonelli, M., Wiebe, N., Culleton, B., House, A., Rabbat, C., Fok, M., McAlister, F., Garg, A., X. (2006). Chronic kidney disease and mortality risk: a systematic review. *Journal of the American Society of. Nephrology*, 17, 7, 2034–47. doi:10.1681/ASN.2005101085. PMID 16738019.
- Van den Haak, M. J., de Jong, M. D. T., & Schellens, P. J. (2003). Retrospective vs. concurrent think aloud protocols: Testing the usability of an online library catalogue. *Behaviour & Information Technology*, 22, 339–351.
- van Oort, L., Schroder, C., & French, D. (2011). What do people think about when they answer the Brief Illness Perception Questionnaire? A 'think-aloud' study. *British Journal of Health Psychology*, 16(2), 231-245.
- Willis, G. B. (2005). *Cognitive interviewing: A tool for improving questionnaire design*. Thousand Oaks, Calif: Sage Publications.
- Willis, G., B., Royston, P., & Bercini, D. (1991). The Use of Verbal Report Methods in the Development and Testing of Survey Questionnaires. *Applied Cognitive Psychology*, 5, 251-267

Paper 3: Critical Appraisal

Word Count: 5,048

Critical Appraisal

Introduction

Accurate measurement of health-related outcomes is central to our understanding of health and illness at both an individual and epidemiological level. Health data informs the development and evaluation of health interventions, as well as health policy and public accounting. McDowell (2006) asserts that subjective reports are the most economically viable evaluative tool to assess health and to provide unique insight into the subjective experience of the respondent. Critically, the data gathered from self-report is only useful to the extent that people make sense of the questions in the manner intended; that is, to the extent to which the scale is measuring the variables it purports to assess.

Cognitive interviewing (also known as think-aloud interviewing or verbal protocols) has enabled researchers to test the assumption, inherent in standardised surveys, that respondents are 1) universally able to understand the questions being asked of them; 2) that this understanding is the understanding intended by the researcher; and 3) that they are subsequently able, and willing, to answer such questions (Collins, 2003). The method has been used to elucidate the way in which respondents understand and interpret questions, and to identify potential problems with questionnaires (Drennan, 2003); including health-related questionnaires (e.g. Horwood, Sutton & Coast, 2013; Heesch, van Uffelen, Hill & Brown, 2010; McCorry, Scullion, McMurray, Houghton & Dempster, 2013). The current thesis focuses on the original cognitive interviewing paradigm: concurrent think-aloud (Ericsson and Simon, 1980, expanded in 1993). In this, respondents are instructed to verbalise their thoughts as they read and complete a questionnaire. Virtually all descriptions of cognitive interviewing include concurrent think-aloud as one possible component (see DeMaio and Landreth 2004; Willis 2005), and some researchers (e.g., Conrad, Blair, and Tracy 2000) have continued to favour the approach.

The current thesis provides a systematic review of the use of think-aloud methods to evaluate questionnaires already in use in the health domain (paper 1) as well as the first evaluation using think-aloud of a questionnaire to measure the use of the emotion regulation strategies of cognitive reappraisal and expressive suppression

(the Emotion Regulation Questionnaire [ERQ; Gross & John, 2003]) in people with End Stage Renal Disease (ESRD) on hospital haemodialysis (paper 2).

The systematic review (paper 1) presents an assessment of the usefulness of think-aloud methods in evaluating health-related questionnaires; including a descriptive account of the aims, samples, methodologies and types of analyses conducted in this area to date. The paper aims to provide an indication firstly, of whether think-aloud has been an effective tool for evaluating questionnaires in a health field. Secondly, the review seeks to outline methodological practices in an attempt to address the lack of consensus within the cognitive interviewing literature (Beatty and Willis, 2007). The paper comments upon the implications of the review for the use and evaluation of survey questionnaires in a health domain. It is hoped that the results will help to guide researchers considering using the think-aloud technique, as well as providing directions for future research.

The empirical study (paper 2) is the first study to examine the ERQ from the perspective of the respondent. The study used think-aloud to examine the processes of interpretation and response to the ERQ with a sample of patients with ESRD in an attempt to determine whether participants made sense of the questionnaire in the way intended by researchers; essentially whether the questionnaire measured what it purports to assess. The paper provides an evaluation of the problems which arose in the task of completing the ERQ and the implications for the validity of the measure. This was a strong study in evaluating the ERQ with an ESRD sample; however, recommendations are made to replicate the study with more representative clinical and non-clinical samples to determine whether or not the problems identified were unique to the population sampled.

These papers will be discussed in this appraisal; with particular emphasis upon the strengths and limitations of the studies, areas for improvement and significant learning points. Firstly, however, it is necessary to mention that the research originated in a somewhat different place. Initially, a different study was designed and pursued to the point of implementation. This was later quite significantly adapted due to the time constraints of the Clinical Psychology Doctoral Programme and a substantial amendment submitted to the Research Ethics Committee. Importantly, it was through the process of switching the focus of the study that some of the main

learning points occurred. In respect of chronology, the appraisal will begin with an account of the original study; its rationale, methodology, and the design processes undertaken. I will then elucidate the course of transition from there to the present thesis.

The Original Study

The original study aimed to examine the determinants of emotional wellbeing in people with ESRD. Patients with ESRD receiving maintenance dialysis suffer from a multitude of physical symptoms, including fatigue, pain, muscle cramps, difficulty with sleep, and sexual dysfunction (Palmer, 2003; Rosas, 2001; Weisbord et al., 2005; Weisbord et al., 2007) and experience substantial impairments in quality of life (QOL) (e.g. Evans et al., 1985; Unruh, Weisbord & Kimmel, 2005). They are further required to undergo regular renal replacement therapy. For those patients on hospital haemodialysis this necessitates their attendance at a specialist Renal Unit on average three times weekly, interfering with employment and social participation. It is unsurprising then, that a large epidemiological study found a near fourfold increase in depression in people with End Stage Renal Disease (ESRD) as compared to healthy individuals (Egede, 2007). Data on the prevalence and impact of anxiety in ESRD are scarcer; however, a review of 55 studies found 38% of people with ESRD experience substantial anxiety (Murtagh, Addington-Hall & Higginson, 2007). Aside from the negative affective and psychosocial impact, a persistent depressive course has been associated with lower perceived health status and quality of life (Cukor, Coplan, Brown, Peterson & Kimmel, 2008), treatment non-adherence (Kimmel et.al., 1998) and increased mortality in an ESRD population (Drayer et al., 2006). Of course, many patients with ESRD do not suffer from depression or anxiety; understandably, there is consequent interest in the cognitive and affective determinants of depression and anxiety in this population. The original study aimed to examine three such potential determinants: illness representations, emotion-regulation strategies and emotional competence.

The Common Sense Model of the Self-Regulation of Health and Illness (CSM; Leventhal, Nerenz & Steele, 1984) suggests that an illness threat elicits cognitive representations (beliefs about the danger of the illness) and emotional representations (emotional states of fear and distress about the illness), collectively called illness representations. Cognitive representations drive procedures or coping strategies to

reduce the health threat, while emotional representations drive strategies to reduce the emotional threat. Importantly, cognitive representations are held to influence emotional representations, and thereby influence emotional coping strategies. Strategies are appraised as to their efficacy in reducing perceived threat and adjusted accordingly. The model is bidirectional, in that new appraisals modify original perceptions of illness threat. Illness perceptions were of interest in the original study as they are associated with a number of physical and psychosocial outcomes in ESRD, including depression (e.g. Chilcot, Wellsted, Davenport & Farrington, 2011).

Cameron and Jago (2008) extended the CSM in line with the process model of emotion regulation (Gross & Thompson, 2007). They suggest that the emotion regulation strategies identified by Gross (1999) may be employed to modify emotional representations of illness. These include i) Attentional deployment (avoid/focus) ii) Proactive behaviour (e.g. treatment adherence) iii) Cognitive change (e.g. reappraisal; Gross, 2002) and iv) Response modulation (e.g. expressive suppression; Gross, 2002). Emotion regulation was of interest in the original study as the strategies are differentially adaptive, for instance, inhibiting the expression of emotion ('suppression') is negatively associated with wellbeing across several medical conditions, including ESRD (Gillanders, Wild, Deighan & Gillanders, 2008). Conversely, modifying the meaning of a situation in a way which reduces emotional threat ('reappraisal') is associated with improved physical and psychological well-being (Karademas, Tsalikou & Tallarou, 2011). This implies that emotion regulation strategies moderate the association between emotional representations and psychological well-being.

The original study also sought to examine emotional competence. Emotional competence refers to a host of processes, skills, and competencies (e.g. attention to feelings ['attention']; clarity about feeling states ['clarity']; attempts to improve mood through optimism ['repair']) that do not directly regulate emotions but make it easier for the individual to behave in socio-emotionally appropriate ways. It has been differentially associated with emotional wellbeing. For instance, attending closely to negative mood states has been shown to increase negative affect and the risk for depression (Scheier & Carver, 1977). Again, this implies that emotional competence would moderate the association between emotional representations and psychological wellbeing.

The original research was designed as a longitudinal study, the primary aim of which was to investigate whether illness representations, emotion-regulation strategies and emotional competence were all independently associated with depression and anxiety (henceforth, collectively referred to as ‘emotional wellbeing’) in people with ESRD. Secondly, it aimed to examine whether emotion regulation strategies and emotional competence moderated the relationship between emotional representations and emotional wellbeing. Further, the study planned to look at the relationships between cognitive and emotional representations and whether emotional representations mediate any association between cognitive representations and emotional well-being.

Transition to the Current Thesis

The original study was planned with the same population as paper 2, however patient measures would have been collected at two time points: all measures at Time 1, then measures of anxiety and depression at Time 2, after 3 months. This required data from 139 participants at time 1 to support the analyses intended. In hindsight this was an ambitious project, requiring the involvement of health professionals for the collection of clinical data and the recruitment of a large number of unwell participants in a busy clinical environment. Essentially, this study transpired to be unfeasible in the timescale.

On reflection, the study design went through a number of iterations, with regard the variables of interest, the measures chosen and the data required to describe the sample; all of which delayed the collection of data. Whilst no doubt an integral part of the research process, I believe the time lost to these challenges was compounded by the fact of experiencing them for the first time.

In October 2013, guidance was sought from the course team as to the feasibility of the study in the time remaining. Following a process of negotiation with all parties involved in the research, it was considered too ambitious for the timescale. A suitable contingency plan was developed and a substantial amendment subsequently submitted to the Local Research Ethics Committee.

Learning from the Process

Changing the study at this juncture was a challenging and stressful experience. The time spent developing the original study was not wasted, however. Through this, I

developed relationships with staff at the Renal Units and with the Research & Department of the target site, facilitating the change. A particular area of learning was around the system for obtaining ethical approval. Given the timeframe, it was important to try and obtain an amendment for the current thesis, rather than to submit an entirely new proposal. This was facilitated by having previously researched the area. In developing the original study I had investigated the literature on determinants of wellbeing in ESRD. This included emotion regulation and an extensive consideration of measures of emotion regulation, leading to the design of the current study (see below). Importantly, I learnt the benefit of strategically matching the tasks required to the time available. Further, perhaps that in some instances, ‘less is more’, in that given the constraints of the training, a smaller, more concise piece of research can be a valuable line of investigation. Most of all it was an opportunity, in a way, for a trial run of the process. Certainly, designing the current study was vastly quicker and simpler the second time around.

Methodological Challenges and Development of the Current Thesis

In developing the original study, the literature was searched for appropriate measures of emotion regulation. Initially, the aim was to measure a wide range of regulation strategies to incorporate the emotion regulation strategies identified by Gross (1999), and incorporated into the CSM by Cameron and Jago (2008). For instance, to include: i) Attentional deployment (avoid/focus) ii) Proactive behaviour (e.g. treatment adherence) iii) Cognitive change (e.g. reappraisal; Gross, 2002) and iv) Response modulation (e.g. expressive suppression; Gross, 2002). Searches did not readily reveal such a comprehensive measure. Attempts to rectify this led to consideration of using two questionnaires: the ERQ and the Brief COPE (Carver, 1997). The ERQ is discussed at length in paper 2; the Brief COPE is an abbreviated version of the original COPE inventory (Carver, Scheier, and Weintraub, 1989). This assesses the use of different coping styles (e.g. self-distraction and venting) in the management of ‘stress’ or negative affect. I also contacted an eminent researcher in the field, James Gross, for advice, and the final decision to use the ERQ was based upon the information he provided. He kindly shared a chapter (John & Eng, 2014) to be included in the new Handbook of Emotion Regulation (2nd Edit. [2014]) which clarifies domains of affect regulation and available measures.

Selection of an appropriate questionnaire was a particular learning point and a strength of the original study. Where there was no comprehensive measure of the different emotion regulation processes, those available had to be assessed as to the suitability and quality for inclusion in the study. John and Eng (2013) highlighted a number of problems with the measures available. The COPE for example, whilst recognised to improve upon previous coping scales (Ways of Coping Questionnaire; Folkman and Lazarus 1980, 1985, 1988), was criticised for being too broad and conceptually ill-defined. For example, several COPE scales (e.g. the Suppression of Competing Activities scale and the Restraint Coping scale) were deemed to reflect problem-focused coping, as opposed to affect regulation. In addition, many of the scales were considered conceptually heterogeneous; making it difficult to establish whether a coping style was adaptive or dysfunctional. This was exemplified by the ‘focus on and venting of emotions’. This was thought to reflect both expressing emotions and being aware of one’s distress. While the former may be considered beneficial in terms of seeking social support, the latter is perhaps more akin to rumination and negatively related to wellbeing. Scales were also defined at different levels of abstraction, from the broad to the specific, and it was unclear how the items were conceptually or causally related. John & Eng (2013, in Gross, 2013) considered the lack of a process, structural, or conceptual model a problem with the measure.

This stood in contrast to the ERQ which was conceptually defined in terms of measuring ‘cognitive reappraisal’ and ‘expressive suppression’. These were also delineated in terms of when during the process of emotional experience they occur. For instance, ‘Cognitive reappraisal’ comprises an attempt to think about a situation in a way which modifies its meaning and reduces emotional threat. This acts before an emotional response has been triggered and was described as an antecedent-focused strategy (Gross, 2001; Gross & John, 2003). ‘Expressive suppression’ meanwhile, is a response-focused strategy, intervening later in the course of the emotional response to minimise or inhibit the outward expression of emotion; that is, once the emotional response is already underway (Gross, 2001; Gross & John, 2003). As discussed in paper 2, Gross and John (2003) found good reliability for the scale; a consistent two-factor structure across samples, ages, and cultures; and evidenced the discriminant validity of the scale. Reliability and validity results have been replicated with translated versions of the scale (e.g. Sala et al., 2012; Balzarotti,

John & Gross, 2010) as well as with children and adolescents (Gullone & Taffe, 2012). It was also possible to discern which of the regulation strategies were adaptive. For instance, reappraisal was associated with more favourable consequences than suppression (Gross, 1998, 2001). This also translated to outcomes in chronic health conditions. For instance, suppression was negatively associated with wellbeing across several medical conditions, including ESRD (Gillanders, Wild, Deighan & Gillanders, 2008). Conversely, reappraisal was associated with improved physical and psychological well-being in cardiac patients (Karademas, Tsalikou & Tallarou, 2011).

On balance, it was decided to measure the more specific and well-defined domains of emotion regulation of: ‘cognitive reappraisal’ and ‘expressive suppression’, using the ERQ. The questionnaire has previously been used with people with ESRD (e.g. Gillanders, 2008), however, the validity has not been investigated with this population. To our knowledge content validity has not been assessed more generally for the ERQ and this was considered a valuable line of enquiry for a pilot study or contingency study. Once the original study was confirmed unfeasible, this contingency plan was undertaken; resulting in the current paper 2.

Paper 2

Transition to the current empirical study has been described above. For continuity therefore, the appraisal will first consider the empirical study, paper 2, before addressing paper 1. Areas of particular importance will be discussed in turn; expanding upon issues presented in the paper itself.

The think-aloud method

Having decided to examine how people make sense of the ERQ, the main consideration was to find an appropriate methodology. Like many of the authors in paper 1, we observed that the think-aloud method had previously been used to achieve similar ends. For instance, the method had been used to elucidate the way in which respondents understand and interpret questions, and to identify potential problems with questionnaires (see Drennan, 2003). A number of studies (Darker & French, 2009; French, Cooke, McLean, Williams & Sutton, 2007; Galansinski, 2008; van Oort, Schroder & French, 2011) have used think-aloud to examine questionnaires already in use in the health domain, for instance, the Theory of

Planned Behaviour Questionnaire (TPB; see Ajzen, 1991, 2002); the Beck Depression Inventory (BDI; Beck, Steer & Brown, 1996); and the Brief Illness Perceptions Questionnaire (Brief IPQ; Broadbent, Petrie, Main, & Weinman, 2006) respectively. Although these scales are used frequently in the field, ‘thinking-aloud’ uncovered a number of problems when participants attempted to complete the questionnaires. The authors of the think-aloud studies subsequently made recommendations to improve the measures, or questioned the overall validity of the questionnaire for the intended purpose.

The benefits and limitations of the think-aloud method have been well described in paper 1 and 2 and the aim of this section is not to repeat these points. The results of the systematic review validate the choice of think-aloud in the empirical study, suggesting that think-aloud is a useful technique to evaluate questionnaires in a health domain. Given the considerations as to the relative efficacy of the various cognitive interviewing techniques, it could be informative to repeat the study using verbal probes or retrospective interviews to compare the processes of interpretation and response to the ERQ.

The Sample

The study is strong if the focus of interest is upon the measurement of emotion regulation in ESRD patients using the ERQ. However, a weakness is in the gender-mix of the sample. The sample comprised 20 men to 5 women (a ratio of 4:1). To some extent this reflects the gender imbalance on the Renal Unit (106 males; 38 females; a ratio of 2.8:1); however, the sample was still under-represented by women. The implications of this were outlined in the discussion section of paper 2. For instance more masculine traits were considered to contribute to some of the views of emotion expressed in the interviews (denial of negative emotion and presenting a positive self-image). Recruitment was fairly challenging on the Units, mainly, in terms of accessing patients at appropriate times and negotiating the research around the essential work of the medical staff. In the timescale it was not possible to recruit a larger sample, or a greater number of women participants. This said, 25 people were recruited and no new themes were emerging in the data at the end of the recruitment period, implying that data saturation had been achieved (Morse, 1995). As discussed in paper 1, 25 is close to the average number of participants recruited (28) in other think-aloud studies of questionnaires. Meaningful

data has also been obtained with far fewer participants (e.g. 6: French et al., 2007; van Oort et al., 2011). Nevertheless, it is conceivable that predominantly older, male participants presented a more stoic self-image to a younger, female interviewer and it would be of interest to replicate the study with a different interviewer. It is also possible that results were influenced by the impacts of ESRD or haemodialysis. For instance participants may have been more cognitively impaired than the general population due to factors such as pain, fatigue or distress. Potentially a non-clinical sample would experience fewer problems with the ERQ, however the current study was particularly interested in the validity of the ERQ with this sample. To conclude therefore, findings may be influenced by the gender-mix of the sample and the clinical population; participants were also mostly white, British and all recruited from one NHS Trust in one part of the UK. All these factors may limit the generalisability of findings to other samples. Results nevertheless point to problematic aspects of the ERQ and future research is needed to replicate this study in more representative and varying samples to inform judgements regarding its overall validity. A strength of the study is that it reveals a need for caution in using the ERQ with people with ESRD, and indicates the need to investigate its use with other populations.

Paper 2: Conclusion

The need for accurate measurement in research is clear, particularly where the findings inform our understanding of concepts such as emotion and could inform interventions to improve health and wellbeing. To this end the current study has sought to explore the processes by which people interpret and respond to the ERQ. In doing so a number of problems have emerged with the questionnaire. The study was thus successful in uncovering issues which would otherwise remain undetected and has contributed to scientific accuracy and rigour. This has simultaneously raised an altogether more human issue; in that one of the authors of the questionnaire, James Gross, previously went out of his way to share information which ultimately facilitated my research. Hence, I am now concerned to repay the consideration, perhaps, by informing him of the research prior to publication. This would maybe be a consideration in conducting a similar study in the future.

Paper 1

Measurement of health-related outcomes informs understanding of health, delivery of interventions and health-care planning, and is frequently undertaken using survey questionnaires. It is vital that researchers can be sure that these instruments measure what they purport to measure. The systematic review (paper 1) presents an assessment of the usefulness of think-aloud methods in evaluating health-related questionnaires; including a descriptive account of the aims, samples, methodologies and types of analyses conducted in this area to date. The paper aimed to provide an indication firstly, of whether think-aloud has been an effective tool for evaluating questionnaires in a health field. Secondly, the review sought to outline methodological practices in an attempt to address the lack of consensus within the cognitive interviewing literature (Beatty and Willis, 2007).

Process Issues

The empirical study informed the development of the research question for the systematic review. It was noted that other studies had successfully used think-aloud methodology to evaluate questionnaires in the health domain. I was interested to know how think-aloud had been used for this purpose; to what extent it had been used; and to what effect. Searches of the literature identified a review of cognitive interviewing (to include think-aloud) (Drennan, 2003), however, this was not specific to think-aloud protocols. Further, the focus was on the pretesting of health surveys, rather than the assessment of questionnaires already in use in the field. A strength of paper 1 is that the review was conducted systematically, rendering the process transparent and replicable. The findings of the review also informed the methodology and analyses chosen for paper 2. It is hoped that the systematic review will similarly inform and encourage future research using this think-aloud.

A strength of the search was that the databases were chosen to represent both psychological and medical literature as indicated by the research question. In addition, the search criteria only specified terms for ‘think-aloud’ and ‘questionnaire’ and did not limit by ‘health’ terms in the search, only in the inclusion criteria. This reduced the risk of excluding potentially relevant papers.

A strength of the current review is that literature was appraised for quality using an amended version of the CASP Qualitative Tool, ensuring that all papers included were of a sufficient standard. To determine the breadth of quality of papers, a scoring

system was developed from 0 (criteria not demonstrated), to 1 (criteria partially demonstrated), to 2 (criteria fully demonstrated). Scores were totalled and compared against pre-determined (determined by the main author in consultation with supervisors) thresholds: low, moderate, high and very high quality. In accordance with guidance on the use of appraisal outcomes (Hannes, 2011; in Noyes, Booth, Hannes, Harden & Harris et al., 2011) it was decided in advance to exclude papers of low quality (a score of <10). No studies were excluded on these grounds, suggesting the existence of good quality research in this area. It is recognised that the application of quality assessment tools is flawed as ratings are subjective in nature. The Cochrane Collaboration (2009) argues that quality rating tools are associated with biased ratings and poor inter-rater reliability. With this in mind, the tool was implemented by the main reviewer and independently rated by a second researcher to establish reliability of quality ratings.

‘Grey’ literature was not included in the current review as the searches were limited to computerised databases. The advantage of only including published papers is that studies will have been peer-reviewed and will have had to demonstrate sufficient quality for publication. Bias exists however, in the type of papers accepted for publication. Positive findings are more likely to be published than negative findings (Button, Ioannidis, Mokrysz, Nosek & Flint 2013); however both are necessary to evaluate the effectiveness of an intervention or experimental technique. A weakness of the current study is that the search strategy potentially mirrored any bias in publication; in that studies which used the think-aloud technique to less effect may not have been published and would not therefore be included in the review. Field and Gillett (2010) recommend contacting experts in the field to request unpublished literature as a way of ameliorating this problem. This was not undertaken due to time constraints, yet presents one way in which the present study could have been improved.

Mixed-Methodology

Critically, the current review sought to elucidate the usefulness of the think-aloud method alone. As discussed in paper 1, this was not always possible as many papers reported findings from an amalgamation of cognitive interviewing techniques. This is probably the main weakness of the current review and was considered the legacy of a lack of methodological clarity within the studies themselves. Results must

therefore be interpreted with caution, as the success of the study in meeting researchers' aims may not be due to think-aloud alone. Recommendations for future research are to differentiate between the findings of each method used. Future research may also wish to compare the relative merits of cognitive interviewing methods used in alone or in combination; in particular, concurrent think-aloud protocols, follow-up probes and retrospective interviews. If it was known which was the most effective technique and under what circumstances, then methodology could be tailored and participant burden reduced.

A weakness of the current review is that it did not comment upon the inherent acceptability of think-aloud interviewing in the evaluation of health-related questionnaires. This was absent from the literature and presents another direction for future investigation.

Key Findings

The review presents a descriptive account of the aims of the studies reviewed; the justification for using think-aloud; populations studied; and methodology. Think-aloud was successfully used to address researchers' aims and was effective at elucidating problems with questionnaires.

A challenge throughout the review was in separating the aims of the review from the aims of the empirical study. That is, in distinguishing aims to assess the usefulness of the think-aloud method in evaluating questionnaires; from aims to evaluate a questionnaire using the think-aloud method. At times, this was a confusing distinction to make; made worse by conducting both at the same time. Indeed, whilst the review purports to assess the usefulness of the think-aloud method in evaluating health-related questionnaires, in fact, the most notable finding may be that questionnaires already commonly in use in the health domain are frequently found to be problematic. This may be the result of using questionnaires as a form of laboratory equipment, ignoring the fact that they are open to interpretation (Mallinson, 2002). Mallinson (2002) considers the correction of misinterpretations, integral in natural conversation, to be missing from the survey interaction; resulting in miscommunication. Future reviews may wish to consider this as an alternative research question. For instance, researchers may choose to systematically review the literature to investigate whether questionnaires already in use in the health domain

actually make sense to respondents in the manner intended by researchers. This may focus upon think-aloud studies, or broaden the methodological criteria to include other methods of assessing content validity, such as verbal probes or structured interviews.

Overall Conclusion

As Mallinson (2002) reports, there is an active industry devoted to the development and application of subjective health measures; however, subjective interpretation is often ignored entirely (Donovan, Frankel & Eyels, 1993). As demonstrated in the studies in paper 1 and the current paper 2, this has implications for the validity of questionnaires. It appears that frequently-used measures are not consistently understood or completed as researchers intended. Results suggest that further research into the validity of health-related questionnaires is warranted; moreover, that this can be effected successfully using think-aloud techniques.

On a more personal level, the experience of completing this research has been a process of both personal and professional growth. From the experience of designing and implementing research to the development and maintenance of professional working relationships, under, at times, trying circumstances; I believe I have become a more competent and confident researcher and Clinical Psychologist; more equipped, I hope, to fulfil the role of scientist-practitioner into the future.

References

- Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, 50, 179–211.
- Ajzen, I. (2002, September). Constructing a TPB questionnaire: Conceptual and methodological considerations. Retrieved from http://chuang.epage.au.edu.tw/ezfiles/168/1168/attach/20/pta_41176_7688352_57138.pdf
- Balzarotti, S., John, O. P., & Gross, J. J. (2010). An Italian adaptation of the Emotion Regulation Questionnaire. *European Journal of Psychological Assessment*, 26, 61-67.
- Beatty, P. C., & Willis, G. B. (2007). Research synthesis: The practice of cognitive interviewing. *Public Opinion Quarterly*, 71, 1-25.
- Beck, A.T., Steer, R.A., & Brown, G.K. (1996). *Beck Depression Inventory II. Manual*. San Antonio, TX: The Psychological Corporation.
- Broadbent, E., Petrie, K. J., Main, J., & Weinman, J. (2006). The Brief Illness Perception Questionnaire. *Journal of Psychosomatic Research*, 60, 631–637
doi:10.1016/j.jpsychores.2005.10.020
- Button, K. S., Ioannidis, J. P. A., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S. J., & Munafò, M. R. (2013). Power failure: Why small sample size undermines the reliability of neuroscience. *Nat Rev Neurosci*, 14(5), 365-376.
- Cameron, L. D. & Jago, L. (2008). Emotion regulation interventions: A common-sense model approach. *British Journal of Health Psychology*, 13, 215-221.
- Carver, C. S. (1997). You want to measure coping but your protocol's too long: Consider the Brief COPE. *International Journal of Behavioral Medicine*, 4, 92-100.
- Carver, C. S., Scheier, M. F., & Weintraub, J. K. (1989). Assessing coping strategies: A theoretically based approach. *Journal of Personality and Social Psychology*, 56, 267–283.
- Chilcot, J., Wellsted, D., Davenport, A. & Farrington, K. (2011) Illness representations and concurrent depression symptoms in haemodialysis patients. *Journal of Health Psychology*, 16, 1127.
- Collins, D. (2003). Pretesting survey instruments: An overview of cognitive methods. *Quality of Life Research*, 12, 229-238.

- Conrad, F., Blair, J., & Tracy, E. (1999). *Verbal reports are data! A theoretical approach to cognitive interviews*. Federal Committee on Statistical Methodology Research Conference, Arlington, USA.
- Cukor, D., Coplan, J., Brown, C., Peterson, R. A., & Kimmel, P. L. (2008). Course of depression and anxiety diagnosis in patients treated with hemodialysis: a 16-month follow-up. *Clin J Am Soc Nephrol*, 3, 6, 1752-1758.
- Darker, C. D., & French, D. P. (2009). What sense do people make of a theory of planned behaviour questionnaire?: a think-aloud study. *Journal of Health Psychology*, 14(7), 861-871.
- DeMaio, T. J., & Landreth, A. (2004). Cognitive interviews: Do different methods produce different results? In S. Presser, J. M. Rothgeb, M. P. Couper, J. T. Lessler, E. Martin, J. Martin, & E. Singer (Eds.), *Methods for testing and evaluating survey questionnaires*. Hoboken, NJ: John Wiley and Sons.
- Donovan, J. L., Frankel, S. J., & Eyels, J. D. (1993). Assessing the need for health status measures. *Journal of Epidemiology Community Health*, 47, 158–162.
- Drayer, R. A., Piraino, B., Reynolds, C. F., 3rd, Houck, P. R., Mazumdar, S., Bernardini, J., et al. (2006). Characteristics of depression in hemodialysis patients: symptoms, quality of life and mortality risk. *Gen Hosp Psychiatry*, 28, 4, 306-312.
- Drennan, J. (2003). Cognitive interviewing: verbal data in the design and pretesting of questionnaires. *J Adv Nurs*, 42:57-63
- Egede, L.E. (2007). Major depression in individuals with chronic medical disorders: Prevalence, correlates and association with health resource utilization, lost productivity and functional disability. *Gen Hosp Psychiatry* 29(5), 409–416.
- Ericsson, K. A. & Simon, H. A. (1980). Verbal reports as data. *Psychological Review*, 87, 215-250
- Ericsson, K. A., & Simon, H. A. (1993). Protocol analysis: Verbal reports as data, revised edition. Cambridge, MA: MIT Press.
- Evans, R. W., Manninen, D. L., Garrison, L.P. Jr., Hart, L. G., Blagg, C. R., Gutman, R. A. ... Lowrie, E. G. (1985). The quality of life of patients with end-stage renal disease. *N Engl J Med* 312, 553–559.
- Field, A.P., & Gillett, R. (2010). How to do a meta-analysis. *British Journal of Mathematical and Statistical Psychology*, 63, 665-694.

- Folkman, S., & Lazarus, R. S. (1980). An analysis of coping in a middle-aged community sample. *Journal of Health and Social Behavior*, 21, 219-239.
- Folkman, S., & Lazarus, R. S. (1985). If it changes it must be a process: Study of emotion and coping during three stages of a college examination. *Journal of Personality and Social Psychology*, 48, 150–170.
- Folkman, S., & Lazarus, R. S. (1988). *Manual for the Ways of Coping Questionnaire*. Palo Alto, CA: Consulting Psychology Press.
- French, D. P., Cooke, R., McLean, N., 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.
- Galasinski, D. (2008). Constructions of the self in interaction with the Beck Depression Inventory. *Health: an Interdisciplinary Journal for the Social Study of Health, Illness & Medicine*, 12(4), 515-533.
- Gillanders, S., Wild, M., Deighan, C., & Gillanders, D. (2008). Emotion regulation, affect, psychosocial functioning, and well-being in hemodialysis patients. *Am J Kidney Dis*. 51, 4, 651-62.
- Gross, J.J. (1998). Antecedent- and response-focused emotion regulation: Divergent consequences for experience, expression, and physiology. *Journal of Personality and Social Psychology*, 74, 224–237.
- Gross, J. J. (1999) Emotion and emotion regulation. In L. A. Pervin (Ed.), *Handbook of personality: Theory and research 2nd ed.* (pp. 525-552). New York, NY: Guilford Press.
- Gross, J.J. (2001). Emotion regulation in adulthood: Timing is everything. *Current Directions In Psychological Science*, 10, 214–219.
- Gross, J. J. (2002). Emotion regulation: Affective, cognitive, and social consequences. *Psychophysiology*, 39, 281-291.
- Gross, J.J., & John, O.P. (2003). Individual differences in two emotion regulation processes: Implications for affect, relationships, and well-being. *Journal of Personality and Social Psychology*, 85, 348–362
- Gross, J.J., & Thompson, R.A. (2007). Emotion regulation: Conceptual foundations. In J.J. Gross (Ed.), *Handbook of emotion regulation*. New York, NY: Guilford Press.

- John, O., P. & Eng, J. S. (2014). Three approaches to individual differences in affect regulation: Conceptualizations, measures . In J. J. Gross (Ed.), *Handbook of emotion regulation (2nd Edit.)*. New York, NY: Guilford Press.
- Gullone & Taffe (2012). The Emotion Regulation Questionnaire for Children and Adolescents (ERQ-CA): A psychometric evaluation. *Psychological Assessment* 24, 2, 409–417. doi: 10.1037/a0025777
- Hannes, K. (2011) Critical appraisal of qualitative research. In: J. Noyes, A. Booth, K. Hannes, A. Harden, J. Harris, S. Lewin, C. Lockwood (Eds.), *Supplementary guidance for inclusion of qualitative research in cochrane systematic reviews of intervention version 1*. Cochrane Collaboration Qualitative Methods Group. Retrieved from <http://cqrng.cochrane.org/supplemental-handbook-guidance>
- Heesch, K. C., van Uffelen, J. G. Z., Hill, R. L., & Brown, W. J. (2010). What do IPAQ questions mean to older adults? Lessons from cognitive interviews. *The International Journal of Behavioral Nutrition and Physical Activity*, 7, 35.
- Horwood, J., Sutton, E., & Coast, J. (2013). Evaluating the face validity of the icecap-o capabilities measure: A "think aloud" study with hip and knee arthroplasty patients. *Applied Research in Quality of Life* Sep(Pagination), No Pagination Specified.
- Karademas, E. C., Tsalikou, C., & Tallarou, M. C. (2011). The impact of emotion regulation and illness-focused coping strategies on the relation of illness-related negative emotions to subjective health. *J Health Psychol* 16, 3, 510-9.
- Kimmel, P. L., Peterson, R. A., Weihs, K. L., Simmens, S. J., Alleyne, S., Cruz, I., & Veis, J. H. (1998). Psychosocial factors, behavioural compliance and survival in urban haemodialysis patients. *Kidney Int*, 54, 1, 245-254.
- Leventhal, H., Nerenz, D., Steele, D. (1984). Illness representations and coping with health threats. In A. Baum, & J. Singer (Eds.), *A handbook of psychology and health: social psychological aspects of health* 4 (pp. 219–52). Hillsdale, NJ: Erlbaum.
- Mallinson, S. (2002). Listening to respondents: a qualitative assessment of the Short-Form 36 Health Status Questionnaire. *Social Science & Medicine*, 54(1), 11–21.
- McDowell, I. (2006). *Measuring health: A guide to rating scales and questionnaires (3rd edit)*. New York, NY: Oxford University Press.

- Morse, J. M. (1995). The significance of saturation. *Qualitative Health Research*, 5, 147-149.
- Murtagh, F. E., Addington-Hall, J., Higginson, I. J. (2007). The prevalence of symptoms in end-stage renal disease: A systematic review. *Adv Chronic Kidney Dis.*, 14, 82-99.
- Palmer, B. F. (2003). Sexual dysfunction in men and women with chronic kidney disease and end-stage kidney disease. *Adv Ren Replace Ther*, 10, 48–60.
- Reitsma, J. B., Rutjes, A. W. S., Whiting, P., Vlassov, V. V., Leeflang, M. M. G., Deeks, J. J. (2009). Assessing methodological quality. In J. J. Deeks, P. M. Bossuyt, C. Gatsonis (Eds.), *Cochrane handbook for systematic reviews of diagnostic test accuracy version 1*. The Cochrane Collaboration, 2009. Available from: <http://srdta.cochrane.org/>
- Rosas, S. E., Joffe, M., Franklin, E., Strom, B.L., Kotzker, W., Brensing, C. ... Feldman, H.I. (2001) Prevalence and determinants of erectile dysfunction in hemodialysis patients. *Kidney Int* 59, 2259–2266
- Sala, M.N., Molina, P., Abler, B., Kessler, H., Vanbrabant, L., & van de Schoot, R. (2012). Measurement invariance of the Emotion Regulation Questionnaire (ERQ). A cross- national validity study. *European Journal of Developmental Psychology*. 9:6, 751-757. doi:10.1080/17405629.2012.690604
- Scheier, M. F., & Carver, C. S. (1977). Self-focused attention and the experience of emotion: Attraction, repulsion, elation, and depression. *Journal of Personality and Social Psychology*, 35, 625–636.
- Unruh, M. L., Weisbord, S.D., Kimmel, P.L. (2005). Health-related quality of life in nephrology research and clinical practice. *Semin Dial*, 18, 82–90
- van Oort, L., Schroder, C., & French, D. (2011). What do people think about when they answer the Brief Illness Perception Questionnaire? A 'think-aloud' study. *British Journal of Health Psychology*, 16(2), 231-245.
- Weisbord, S.D., Fried, L.F., Arnold, R.M., Fine, M.J., Levenson, D.J., Peterson, R.A., Switzer, G.E. (2005). Prevalence, severity, and importance of physical and emotional symptoms in chronic hemodialysis patients. *J Am Soc Nephrol* 16, 2487–2894.
- Weisbord, S.D., Fried, L.F., Mor, M.K., Resnick, A.L., Unruh, M.L., Palevsky, P.M., Levenson, D.J., Cooksey, S.H., Fine, M.J., Kimmel, P.L., Arnold, R.M.

(2007). Renal provider recognition of symptoms in patients on maintenance hemodialysis. *Clin J Am Soc Nephrol* 2, 960–967.

Willis, G. B. (2005). *Cognitive interviewing: A tool for improving questionnaire design*. Thousand Oaks, Calif: Sage Publications.

Appendices

Appendix A: British Journal of Health Psychology: Author Guidelines

Author Guidelines

The aim of the British Journal of Health Psychology is to provide a forum for high quality research relating to health and illness. The scope of the journal includes all areas of health psychology across the life span, ranging from experimental and clinical research on aetiology and the management of acute and chronic illness, responses to ill-health, screening and medical procedures, to research on health behaviour and psychological aspects of prevention. Research carried out at the individual, group and community levels is welcome, and submissions concerning clinical applications and interventions are particularly encouraged.

The types of paper invited are:

- papers reporting original empirical investigations, using either quantitative or qualitative methods;
- theoretical papers which may be analyses or commentaries on established theories in health psychology, or presentations of theoretical innovations;
- review papers, which should aim to provide systematic overviews, evaluations and interpretations of research in a given field of health psychology; and
- methodological papers dealing with methodological issues of particular relevance to health psychology.

1. Circulation

The circulation of the Journal is worldwide. Papers are invited and encouraged from authors throughout the world.

2. Length

Papers should normally be no more than 5000 words (excluding the abstract, reference list, tables and figures), although the Editor retains discretion to publish papers beyond this length in cases where the clear and concise expression of the scientific content requires greater length.

3. Editorial policy

The Journal receives a large volume of papers to review each year, and in order to make the process as efficient as possible for authors and editors alike, all papers are initially examined by the Editors to ascertain whether the article is suitable for full peer review. In order to qualify for full review, papers must meet the following criteria:

- the content of the paper falls within the scope of the Journal
- the methods and/or sample size are appropriate for the questions being addressed
- research with student populations is appropriately justified
- the word count is within the stated limit for the Journal (i.e. 5000 words)

4. Submission and reviewing

All manuscripts must be submitted via [Editorial Manager](#). You may like to use the [Submission Checklist](#) to help you prepare your manuscript. The Journal operates a policy of anonymous peer review. Authors must suggest three reviewers when submitting their manuscript, who may or may not be approached by the Associate Editor dealing with the paper. Before submitting, please read the [terms and conditions of submission](#) and the [declaration of competing interests](#).

5. Manuscript requirements

- Contributions must be typed in double spacing with wide margins. All sheets must be numbered.
- Manuscripts should be preceded by a title page which includes a full list of authors and their affiliations, as well as the corresponding author's contact details. A template can be downloaded from [here](#).
- Statement of Contribution: All authors are required to provide a clear summary of 'what is already known on this subject?' and 'what does this study add?'. Authors should identify existing research knowledge relating to the specific research question and give a summary of the new knowledge added by your study. Under each of these headings, please provide 2-3 (maximum) clear outcome statements (not process statements of what the paper does); the statements for 'what does this study add?' should be presented as bullet points of no more than 100 characters each. The Statement of Contribution should be a separate file.
- Tables should be typed in double spacing, each on a separate page with a self-explanatory title. Tables should be comprehensible without reference to the text. They should be placed at the end of the manuscript with their approximate locations indicated in the text.
- Figures can be included at the end of the document or attached as separate files, carefully labelled in initial capital/lower case lettering with symbols in a form consistent with text use. Unnecessary background patterns, lines and shading should be avoided. Captions should be listed on a separate sheet. The resolution of digital images must be at least 300 dpi.
- For articles containing original scientific research, a structured abstract of up to 250 words should be included with the headings: Objectives, Design, Methods, Results, Conclusions. Review articles should use these headings: Purpose, Methods, Results, Conclusions.
- For reference citations, please use APA style. Particular care should be taken to ensure that references are accurate and complete. Give all journal titles in full and provide doi numbers where possible for journal articles. For example:

Author, A., Author, B., & Author, C. (1995). *Title of book*. City, Country: Publisher.
Author, A. (2013). Title of journal article. *Name of journal*, 1, 1-16. doi: 10.1111/bjep.12031
- SI units must be used for all measurements, rounded off to practical values if appropriate, with the imperial equivalent in parentheses.
- In normal circumstances, effect size should be incorporated.
- Authors are requested to avoid the use of sexist language.

- Authors are responsible for acquiring written permission to publish lengthy quotations, illustrations, etc. for which they do not own copyright. For guidelines on editorial style, please consult the [APA Publication Manual](#) published by the American Psychological Association.
- Manuscripts describing clinical trials are encouraged to submit in accordance with the [CONSORT statement](#) on reporting randomised controlled trials.

6. Supporting information

Supporting Information can be a useful way for an author to include important but ancillary information with the online version of an article. Examples of Supporting Information include appendices, additional tables, data sets, figures, movie files, audio clips, and other related nonessential multimedia files. Supporting Information should be cited within the article text, and a descriptive legend should be included. Please indicate clearly on submission which material is for online only publication. It is published as supplied by the author, and a proof is not made available prior to publication; for these reasons, authors should provide any Supporting Information in the desired final format.

For further information on recommended file types and requirements for submission, please visit the [Supporting Information page](#) on Author Services.

7. OnlineOpen

OnlineOpen is available to authors of primary research articles who wish to make their article available to non-subscribers on publication, or whose funding agency requires grantees to archive the final version of their article. With OnlineOpen, the author, the author's funding agency, or the author's institution pays a fee to ensure that the article is made available to non-subscribers upon publication via Wiley Online Library, as well as deposited in the funding agency's preferred archive. A full list of [terms and conditions](#) is available on Wiley Online Library.

Any authors wishing to send their paper OnlineOpen will be required to complete the [payment form](#).

Prior to acceptance there is no requirement to inform an Editorial Office that you intend to publish your paper OnlineOpen if you do not wish to. All OnlineOpen articles are treated in the same way as any other article. They go through the journal's standard peer-review process and will be accepted or rejected based on their own merit.

8. Author Services

Author Services enables authors to track their article – once it has been accepted – through the production process to publication online and in print. Authors can check the status of their articles online and choose to receive automated e-mails at key stages of production. The author will receive an e-mail with a unique link that enables them to register and have their article automatically added to the system. Please ensure that a complete e-mail address is provided when submitting the manuscript. Visit [Author Services](#) for more details on online production tracking and for a wealth of resources including FAQs and tips on article preparation, submission and more.

9. Copyright and licences

If your paper is accepted, the author identified as the formal corresponding author for the paper will receive an email prompting them to login into Author Services, where via the Wiley Author Licensing Service (WALS) they will be able to complete the licence agreement on behalf of all authors on the paper.

For authors signing the copyright transfer agreement

If the OnlineOpen option is not selected the corresponding author will be presented with the copyright transfer agreement (CTA) to sign. The terms and conditions of the CTA can be previewed in the samples associated with the [Copyright FAQs](#).

For authors choosing OnlineOpen

If the OnlineOpen option is selected the corresponding author will have a choice of the following Creative Commons Licence Open Access Agreements (OAA):

- Creative Commons Attribution Non-Commercial Licence (CC-BY-NC)
- Creative Commons Attribution Non-Commercial -NoDerivs Licence (CC-BY-NC-ND)

To preview the terms and conditions of these open access agreements please visit the [Copyright FAQs](#) and you may also like to visit the [Wiley Open Access Copyright and Licence](#) page.

If you select the OnlineOpen option and your research is funded by The Wellcome Trust and members of the Research Councils UK (RCUK) you will be given the opportunity to publish your article under a CC-BY licence supporting you in complying with Wellcome Trust and Research Councils UK requirements. For more information on this policy and the Journal's compliant self-archiving policy please visit our [Funder Policy](#) page.

10. Colour illustrations

Colour illustrations can be accepted for publication online. These would be reproduced in greyscale in the print version. If authors would like these figures to be reproduced in colour in print at their expense they should request this by completing a [Colour Work Agreement](#) form upon acceptance of the paper.

11. Pre-submission English-language editing

Authors for whom English is a second language may choose to have their manuscript professionally edited before submission to improve the English. A list of independent suppliers of editing services can be found in [Author Services](#). All services are paid for and arranged by the author, and use of one of these services does not guarantee acceptance or preference for publication.

12. The Later Stages

The corresponding author will receive an email alert containing a link to a web site. A working e-mail address must therefore be provided for the corresponding author. The proof can be downloaded as a PDF (portable document format) file from this site. Acrobat Reader will be required in order to read this file. This software can be downloaded (free of charge)

from [Adobe's web site](#). This will enable the file to be opened, read on screen and annotated direct in the PDF. Corrections can also be supplied by hard copy if preferred. Further instructions will be sent with the proof. Hard copy proofs will be posted if no e-mail address is available. Excessive changes made by the author in the proofs, excluding typesetting errors, will be charged separately.

13. Early View

British Journal of Health Psychology is covered by the Early View service on Wiley Online Library. Early View articles are complete full-text articles published online in advance of their publication in a printed issue. Articles are therefore available as soon as they are ready, rather than having to wait for the next scheduled print issue. Early View articles are complete and final. They have been fully reviewed, revised and edited for publication, and the authors' final corrections have been incorporated. Because they are in final form, no changes can be made after online publication. The nature of Early View articles means that they do not yet have volume, issue or page numbers, so they cannot be cited in the traditional way. They are cited using their Digital Object Identifier (DOI) with no volume and issue or pagination information. Eg Jones, A.B. (2010). Human rights Issues. *Journal of Human Rights*. Advance online publication. doi:10.1111/j.1467-9299.2010.00300.x

**Appendix B: Confirmation of Ethical Approval for the original study from
NRES Committee North West - Liverpool East**

NRES Committee North West - Liverpool East

HRA NRES Centre Manchester
Barlow House
3rd Floor
4 Minshull Street
Manchester
M1 3DZ

Telephone: 0161 625 7832
Facsimile: 0161 625 7299

26 September 2013

Miss Anna Phillips
Trainee Clinical Psychologist
Manchester Mental Health and Social Care Trust
Department of Clinical Psychology
2nd Floor, Zochonis Building
Brunswick Street
M13 9PL

Dear Miss Phillips

Study title:	Illness Representations, emotion regulation strategies and emotional wellbeing in people with End Stage Renal Disease.
REC reference:	13/NW/0645
Protocol number:	n/a
IRAS project ID:	121467

The Research Ethics Committee reviewed the above application at the meeting held on 19 September 2013. Thank you for attending to discuss the application.

We plan to publish your research summary wording for the above study on the NRES website, together with your contact details, unless you expressly withhold permission to do so. Publication will be no earlier than three months from the date of this favourable opinion letter. Should you wish to provide a substitute contact point, require further information, or wish to withhold permission to publish, please contact the Co-ordinator Miss Helen Penistone, nrescommittee.northwest-liverpooleast@nhs.net.

Ethical opinion

The Committee explained that they had read the application thoroughly and found it to be very well written. The Committee referred to the *Causes of My Illness* questionnaire and asked whether it would be possible to add another possible cause to the table. The Committee were particularly thinking of the possibility that the renal disease was caused by prescribed medications. You agreed that this was a very valid suggestion but explained that you were unsure how this could be done. You could contact the author.

The Committee understood that the questionnaire was validated but advised that sometimes authors are happy for modifications to be made so that the questionnaire is applicable to a specific illness. The Committee suggested that one of the current possible causes listed could be omitted so that this could be added. It is important to consider how the responses would be

A Research Ethics Committee established by the Health Research Authority

analysed before making a change. The Committee also suggested that when contacting the author you could explain that the request had been made by the REC.

The Committee referred to the Consent Form and explained that a point should be added to request consent to inform the participant's GP of their involvement in the study. You advised that you would not be contacting the GP but would contact the renal team and this was already included in the Consent Form.

The members of the Committee present gave a favourable ethical opinion of the above research on the basis described in the application form, protocol and supporting documentation, subject to the conditions specified below.

The Committee **suggested** enquiring about whether it would be possible to alter the *Causes of My Illness* questionnaire to add another possible cause of 'other'.

Ethical review of research sites

NHS Sites

The favourable opinion applies to all NHS sites taking part in the study, subject to management permission being obtained from the NHS/HSC R&D office prior to the start of the study (see "Conditions of the favourable opinion" below).

Conditions of the favourable opinion

The favourable opinion is subject to the following conditions being met prior to the start of the study.

Management permission or approval must be obtained from each host organisation prior to the start of the study at the site concerned.

Management permission ("R&D approval") should be sought from all NHS organisations involved in the study in accordance with NHS research governance arrangements.

Guidance on applying for NHS permission for research is available in the Integrated Research Application System or at <http://www.rdforum.nhs.uk>.

Where a NHS organisation's role in the study is limited to identifying and referring potential participants to research sites ("participant identification centre"), guidance should be sought from the R&D office on the information it requires to give permission for this activity.

For non-NHS sites, site management permission should be obtained in accordance with the procedures of the relevant host organisation.

Sponsors are not required to notify the Committee of approvals from host organisations

It is responsibility of the sponsor to ensure that all the conditions are complied with before the start of the study or its initiation at a particular site (as applicable).

Approved documents

The documents reviewed and approved at the meeting were:

<i>Document</i>	<i>Version</i>	<i>Date</i>
REC application 121467/492020/1/267		13 August 2013
Protocol	2	14 August 2013
Investigator CV - Miss Anna Phillips		17 July 2013

A Research Ethics Committee established by the Health Research Authority

Investigator CV - Alison Wearden		03 January 2013
Investigator CV - David French		01 February 2013
Participant Information Sheet	2	14 August 2013
Participant Consent Form	1	17 July 2013
GP/Consultant Information Sheets	1	17 July 2013
Letter from Sponsor	from Lynne MacRae	06 August 2013
Questionnaire: Hospital Anxiety and Depression Scale (HADS)		
Questionnaire: Trait Meta-Mood Scale (TMMS)		
Questionnaire: Illness Perception Questionnaire (IPQ-R)		
Questionnaire: Emotion Regulation Questionnaire (ERQ)	1	17 July 2013
Questionnaire: Demographic Questionnaire	1	17 July 2013
Other: Distress Protocol	1	17 July 2013
Referees or other scientific critique report	Letter from Dr Dougal Hare	29 November 2012
Evidence of insurance or indemnity	Letter from Lynne MacRae	06 August 2013

Membership of the Committee

The members of the Ethics Committee who were present at the meeting are listed on the attached sheet.

Statement of compliance

The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.

After ethical review

Reporting requirements

The attached document "After ethical review – guidance for researchers" gives detailed guidance on reporting requirements for studies with a favourable opinion, including:

- Notifying substantial amendments
- Adding new sites and investigators
- Notification of serious breaches of the protocol
- Progress and safety reports
- Notifying the end of the study

The NRES website also provides guidance on these topics, which is updated in the light of changes in reporting requirements or procedures.

Feedback

You are invited to give your view of the service that you have received from the National Research Ethics Service and the application procedure. If you wish to make your views known please use the feedback form available on the website.

Further information is available at National Research Ethics Service website > After Review

13/NW/0645	Please quote this number on all correspondence
-------------------	---

A Research Ethics Committee established by the Health Research Authority

We are pleased to welcome researchers and R & D staff at our NRES committee members' training days – see details at <http://www.hra.nhs.uk/hra-training/>

With the Committee's best wishes for the success of this project.

Yours sincerely



On behalf of
Mrs Glenys J Hunt
Chair

Email: nrescommittee.northwest-liverpooleast@nhs.net

Enclosures: List of names and professions of members who were present at the meeting and those who submitted written comments

"After ethical review – guidance for researchers"

Copy to: Ms Lynne Macrae
The University of Manchester

Mr Iain McLean
Central Manchester University Hospitals NHS Foundation Trust

NRES Committee North West - Liverpool East

Attendance at Committee meeting on 19 September 2013

Committee Members:

<i>Name</i>	<i>Profession</i>	<i>Present</i>	<i>Notes</i>
Mr John Bridson	Clinical Ethicist	Yes	
Dr Zoe Edwards	Clinical Psychologist	Yes	
Mrs Elizabeth Gordon		No	
Mrs Maureen Hendry	Pharmacist	Yes	
Mrs Glenys J Hunt	Lay member	Yes	
Mr Chris Irving	Biomedical Scientist	Yes	
Dr S.M. Mostafa		Yes	
Professor Ebrahim Khalil Naderali	Professor of Human Physiology	Yes	
Mr Alex Newgrosh	Quality Assurance Manager	Yes	
Professor Neil Pender	Professor of Orthodontics	Yes	
Mrs Jean Pownceby	Lay Member	Yes	
Miss Kimberley Saint	Trainee Clinical Scientist	No	
Dr Richard Sarginson	Consultant (Anaesthesia/PICU)	Yes	
Dr Peter Walton	Lay Member	Yes	

**Appendix C: Confirmation of Ethical Approval for the Substantial Amendment
from NRES Committee North West - Liverpool East**



Health Research Authority

National Research Ethics Service

NRES Committee North West - Liverpool East

HRA NRES Centre Manchester
Barlow House
3rd Floor
4 Minshull Street
Manchester
M1 3DZ

Tel: 0161 625 7832
Fax: 0161 625 7299

18 December 2013

Miss Anna Phillips
Trainee Clinical Psychologist
Manchester Mental Health and Social Care Trust
Department of Clinical Psychology
2nd Floor, Zochonis Building
Brunswick Street
M13 9PL

Dear Miss Phillips

Study title:	Illness Representations, emotion regulation strategies and emotional wellbeing in people with End Stage Renal Disease.
REC reference:	13/NW/0645
Protocol number:	n/a
Amendment number:	Substantial Amendment 1
Amendment date:	14 August 2013
IRAS project ID:	121467

The above amendment was reviewed by the Sub-Committee in correspondence.

Ethical opinion

Approval was sought for preliminary assessment of the validity of the Emotion Regulation Questionnaire.

The Committee queried what the time frame for transcription and subsequent destruction of audio recordings would be. You revised the protocol to clarify this point.

The members of the Committee taking part in the review gave a favourable ethical opinion of the amendment on the basis described in the notice of amendment form and supporting documentation.

Approved documents

The documents reviewed and approved at the meeting were:

Document	Version	Date
Notice of Substantial Amendment (non-CTIMPs): Substantial Amendment 1		14 August 2013
Protocol	3 Amendment 1	26 November 2013
Changes to the IRAS Form	Amendment 1	22 November 2013
Changes to the Protocol	Amendment 1	22 November 2013
GP/Consultant Information Sheets	1	22 November 2013
Participant Information Sheet: for Stage 1 of the study	1	22 November 2013
Participant Consent Form: for Stage 1 of the study	1	22 November 2013

Membership of the Committee

The members of the Committee who took part in the review are listed on the attached sheet.

R&D approval

All investigators and research collaborators in the NHS should notify the R&D office for the relevant NHS care organisation of this amendment and check whether it affects R&D approval of the research.

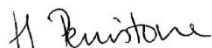
Statement of compliance

The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.

We are pleased to welcome researchers and R & D staff at our NRES committee members' training days – see details at <http://www.hra.nhs.uk/hra-training/>

13/NW/0645:	Please quote this number on all correspondence
--------------------	---

Yours sincerely



On behalf of
Mrs Glenys J Hunt
Chair

E-mail: nrescommittee.northwest-liverpooleast@nhs.net

Enclosures: List of names and professions of members who took part in the review

Copy to: Mr Iain McLean,
Central Manchester University Hospitals NHS Foundation Trust

Ms Lynne Macrae,
University of Manchester

A Research Ethics Committee established by the Health Research Authority

NRES Committee North West - Liverpool East

Attendance at Sub-Committee of the REC meeting

<i>Name</i>	<i>Profession</i>	<i>Capacity</i>
Mrs Glenys J Hunt	Lay member	Lay
Professor Neil Pender	Professor of Orthodontics	Expert

Also in attendance:

<i>Name</i>	<i>Position (or reason for attending)</i>
Miss Helen Penistone	REC Manager

Appendix D: R&D Letter of Approval of Substantial Amendment

Research & Development
1st Floor Post Graduate Centre
Manchester Royal Infirmary
Oxford Road
Manchester M13 9WL
Tel: 0161-276-3340
Fax: 0161-276-5766
lorraine.broadfoot@cmft.nhs.uk

Ref: R03383-LTR 13-Phillips-Substantial Amendment 1

Carla Barrett
Renal Clinical Trials Manager
Renal Transplant Admin Offices
5th Floor, St Mary's
Central Manchester University Hospitals NHS Foundation Trust
Oxford Road
Manchester
M13 9WL

Dear Carla,

Re: R03383: Illness Representations, emotion regulation strategies and emotional wellbeing in people with End Stage Renal Disease.
REC Reference: 13/NW/0645
Principal Investigator: Anna Phillips
Amendment Number: Substantial Amendment 1
Amendment Date: 14 August 2013

Thank you for your correspondence informing the department of an amendment to the above project; we acknowledge receipt of the following and approve the amendment.

Document	Version Number	Dated
NRES Amendment Approval Letter	Substantial Amendment 1	18 December 2013
Notice of Substantial Amendment (non-CTIMPs): Substantial Amendment 1		14 August 2013
Protocol	3, Amendment 1	26 November 2013
Changes to the IRAS Form	Amendment 1	22 November 2013
Changes to Protocol	Amendment 1	22 November 2013
GP/Consultant Information Sheets	1	22 November 2013
Participant Information Sheet: For Stage 1 of the Study	1	22 November 2013
Participant Consent Form: For Stage 1 of the Study	1	22 November 2013

We have amended the Trust's database to reflect these changes as required.

I would like to take this opportunity to thank you for keeping the Trust informed and wish you continued success with your project.

Yours sincerely



Lorraine Broadfoot
Research Operations Manager

Date: 16/01/2014

cc Anna Phillips
Iain McLean

Appendix E: The Emotion Regulation Questionnaire (ERQ; Gross & John, 2003)

Emotion Regulation Questionnaire (ERQ)

We would like to ask you some questions about your emotional life, in particular, how you control (that is, regulate and manage) your emotions.

The questions below involve two distinct aspects of your emotional life. One is your emotional experience, or what you feel like inside.

The other is your emotional expression, or how you show your emotions in the way you talk, gesture, or behave.

Although some of the following questions may seem similar to one another, they differ in important ways. For each item, please answer using the following scale:

1-----2-----3-----4-----5-----6-----7
Strongly Neutral Strongly
Disagree Agree

1. ____ When I want to feel more *positive* emotion (such as joy or amusement), I *change what I'm thinking about*.
2. ____ I keep my emotions to myself.
3. ____ When I want to feel less *negative* emotion (such as sadness or anger), I *change what I'm thinking about*.
4. ____ When I am feeling *positive* emotions, I am careful not to express them.
5. ____ When I'm faced with a stressful situation, I make myself *think about it* in a way that helps me stay calm.
6. ____ I control my emotions by *not expressing them*.
7. ____ When I want to feel more *positive* emotion, I *change the way I'm thinking* about the situation.
8. ____ I control my emotions by *changing the way I think* about the situation I'm in.
9. ____ When I am feeling *negative* emotions, I make sure not to express them.
10. ____ When I want to feel less *negative* emotion, I *change the way I'm thinking* about the situation.

Appendix F: Demographic Questionnaire

Demographic Questionnaire

This information collects information about yourself and your background. Please read and answer every question. All information provided will be treated in confidence and will not be made available to any other source without your approval.

1. Your Family

What is your current relationship status?

- ☐ Single
☐ Married/living with partner
☐ Separated or divorced
☐ Widowed

2. Caring Responsibilities

Do you have caring responsibilities for anybody else (e.g. children/ older people)?

- ☐ Yes
☐ No

3. Your ethnicity

With which ethnic group do you identify?

White

- A) British
B) Irish
C) Any other white background

Mixed

- D) White and black Caribbean
E) White and black African
F) White and Asian
G) Any other mixed background

Asian or Asian British

- H) Indian
J) Pakistani
K) Bangladeshi
L) Any other Asian Background

Black or Black British

- M) Caribbean
N) African
P) Any other black background

Other ethnic groups

- R) Chinese
S) Any other ethnic group

4. Education

What is your highest level of education?

- ☐ No qualifications

Version number:1

Date: 17/07/2013

- ☐ GCSEs, CSEs, or O-levels
- ☐ A levels/ BTEC
- ☐ Trade/apprenticeship
- ☐ University degree
- ☐ Other (please specify) _____

To end of year ____

5. Employment Status

- ☐ Full time
- ☐ Part time
- ☐ Home duties
- ☐ Unemployed

Version number:1

Date: 17/07/2013

Appendix G: Participant Information Sheet

**Project: What sense do people with Chronic Kidney Disease make of the
Emotion Regulation Questionnaire?**

You are invited to take part in a research study. Before you decide if you want to take part, it is important for you to understand why the research is being done and what it will involve. Please take time to read or listen to the following information carefully and decide if you wish to take part.

Feel free to ask any questions and talk to others before you decide whether to take part.

Thank you for reading this.

What is the study about?

We will shortly be beginning a study looking at how people manage the emotions brought up by their kidney disease. For this study, we will use a questionnaire about how people regulate their emotions. We want to check that people understand the questions in the way that we meant them.

Who will conduct the research?

The research is a 'Think-Aloud' study conducted by Anna Phillips, a doctorate student on a Clinical Psychology course at the University of Manchester who is being supervised by Professor Alison Wearden and Professor David French.

Why is the research being done?

It is important to know whether the people who complete the questionnaires make sense of the questions in the same way as the researchers intended. This is especially important if the results of the questionnaires are used to design interventions to help people regulate their emotions.

Who can take part in the study?

If you have Chronic Kidney Disease and you are aged between 18 and 80 years of age, and you attend the Renal Unit for haemodialysis you are invited to take part in the study. You

Participant Information Sheet for stage 1 of the study (i.e. Validity check of the Emotion Regulation Questionnaire). Version number 1 Amendment 1

Date: 22/11/2013

would also need to be able to understand written or spoken English to be able to fill in the questionnaire and to talk aloud while filling it in.

What does the research involve?

You will be asked to complete the Emotion Regulation Questionnaire whilst thinking aloud. What I mean by 'think aloud' is that I want you to tell me everything you are thinking as you read each question and decide how to answer it. This will take around 15 minutes in total. You will be audio-recorded whilst you do this and all recordings will be transcribed.

The researcher will also look at your patient notes to find information about your kidney disease and make a note of any other health conditions you have.

How will the information be kept confidential?

All person identifiable information will be kept anonymous and strictly confidential. This means that your identity will be kept private at all times. Tracey Hepburn, Administrator at The University of Manchester will transcribe anonymised audio-recordings of your interview. This will be conducted on secure computers at The University of Manchester from encrypted data drives. Professor Alison Wearden and Professor David French may see anonymised transcripts of your interviews so they will not be able to link the transcript to the participant. The anonymous data will be stored on confidential University computers and hard-copies will be in a locked filing cabinet at the University of Manchester, Zochonis building. All data will be destroyed confidentially after 5 years. Any personal contact details will be also be destroyed confidentially after 5 years.

Your personal data may be accessed by individuals responsible for auditing and monitoring the conduct of the study on behalf of the University, NHS or regulatory authorities who will have access for this purpose. This is routine for research.

All data will be anonymised before it is published so that no one person can be identified by any of the data when it is made public (either as part of my doctorate thesis, in academic journal articles or as part of any conference publications).

What are the benefits of taking part in this study?

There are no direct benefits from taking part in the research. The full study will add to the understanding of why some people with Chronic Kidney Disease experience more anxiety or depression. To do this we first need to know whether the Emotion Regulation Questionnaire is measuring what it is designed to measure. This knowledge could in future be used to design interventions to help people to cope better with difficult emotions brought up by their kidney disease in order to promote better emotional wellbeing.

What are the risks of taking part in this study?

The only risks are that you may feel more burdened by taking part, or that you may be upset by your answers to some questions.

What if I get upset by the questions asked?

If any of your answers to the questions upset or worry you, please inform Anna Phillips and she will ask if you would like to speak to someone about what is distressing you. If you are Participant Information Sheet for stage 1 of the study (i.e. Validity check of the Emotion Regulation Questionnaire). Version number 1 Amendment 1

Date: 22/11/2013

upset by an aspect of your kidney disease or treatment, then she will ask if you would like her to inform the Renal Team. Alternatively, with your permission, she could refer you to the Clinical Psychologist at the Renal Unit.

If you feel distressed after you have completed the questionnaires, please contact the Renal Team.

It will be your choice whether you would like Anna to tell anyone else about your concerns. The only exceptions to this are if she is worried that you could be at risk of harm, or that someone else was at risk of harm. In these circumstances she may have to tell someone else even if you do not wish her to.

What if I do not want to take part any more?

Taking part is your choice. You do not have to take part and can stop at any time. You do not have to give a reason for changing your mind. If you stop doing the study it will not affect your care or future treatment on the renal unit. It will not affect your legal rights. If you decide to stop taking part in the study you will be asked whether the study can keep the information you have already given. If you decide no, this information will be removed from the study.

What will happen to the results?

It is the intention of the researcher that this research will be written up and published in relevant journals that are read by health professionals. No identifiable details will be published. For example, we would never use your name, contact details or date of birth. Anonymised quotes may be used in the publication. The results of the final analysis will be fed back to interested staff and service-users from the Manchester Royal Infirmary as this is the site from which participants have been recruited. However, you will never be identified.

What if I want to make a complaint?

If you have a concern about any aspect of the study you can contact Anna Phillips via email: anna.phillips-2@postgrad.manchester.ac.uk, or telephone: 0161 306 0400. If she is unable to help you or if you wish to make a complaint about the study please contact the Research Practice and Governance Co-ordinator by either writing to 'The Research Practice and Governance Co-ordinator, Research Office, Christie Building, The University of Manchester, Oxford Road, Manchester M13 9PL', by emailing: Research.complaints@manchester.ac.uk, or by telephoning 0161 275 7583 or 275 8093.

Compensation

In the unlikely event that something does go wrong and you are harmed during the research you may have grounds for a legal action for compensation against the University of Manchester or Central Manchester NHS Foundation Trust but you may have to pay your legal costs. The normal National Health Service complaints mechanisms will still be available to you.

Who has reviewed this study?

Participant Information Sheet for stage 1 of the study (i.e. Validity check of the Emotion Regulation Questionnaire). Version number 1 Amendment 1

Date: 22/11/2013

This study has been reviewed and approved by an NHS Ethics Committee (North West Liverpool East).

Contact for further information

If you have any questions about the study contact Anna Phillips on anna.phillips-2@postgrad.manchester.ac.uk or telephone: 0161 306 0400.

Participant Information Sheet for stage 1 of the study (i.e. Validity check of the Emotion Regulation Questionnaire). Version number 1 Amendment 1

Date: 22/11/2013

Appendix H: Participant Consent Form

Consent Form

Project: What sense do people with Chronic Kidney Disease make of the Emotion Regulation Questionnaire?

Researchers: Anna Phillips (Trainee Clinical Psychologist), Professor Alison Wearden and Professor David French.

Chief Investigator: Anna Phillips, Trainee Clinical Psychologist

Contact details: The Academic Division of Clinical Psychology
2nd Floor, Zochonis Building
University of Manchester
Brunswick Street
Manchester M13 9PL
Email: anna.phillips-2@postgrad.manchester.ac.uk
Tel: 0161 306 0400

Please initial box

1. I have read and understood the information sheet.
Anything I didn't understand has been talked about. Any
questions I had have been answered. ☐
2. I agree to take part in the above study. The study includes:
 - Filling in a questionnaire whilst thinking aloud ☐
3. I understand that taking part is my choice. I don't have to take
part and I can stop at any time. I understand that I do not have
to give a reason for changing my mind. If I stop doing the study,
it will not affect my care. It will not affect my legal rights. ☐

Participant Consent form for stage 1 of the study (i.e. Validity check of the Emotion Regulation Questionnaire). Version 1; Amendment 1

Date: 22/11/2013

4. I understand that Anna will look at my patient records to find relevant information about my kidney disease and make a note of any other health conditions I have. ☐
5. I understand that Anna will tell the Renal Team that I am taking part. I understand that any information I give will be kept safe. It will not be shared with anyone, unless Anna thinks that I am at risk or she thinks that someone else is at risk. Also, reports and publications will not use my real name. No one will know who I am from the information in reports. ☐
6. I understand and consent to my interview being audio-recorded. I understand that this recording will be anonymised and then transcribed. ☐
7. I understand that direct, anonymised quotes from my interview may be used to illustrate findings in reports or publications ☐
8. I understand that sections of data collected during the study may be looked at by responsible individuals from the University of Manchester, from regulatory authorities or from the NHS Trust, where it is relevant to my taking part in the study. I give permission for these people to have access to this data. ☐

I agree to take part in the study. ☐

_____	_____	_____
Name of participant	Date	Signature

_____	_____	_____
Name of participant	Date	Signature

Participant Consent form for stage 1 of the study (i.e. Validity check of the Emotion Regulation Questionnaire). Version 1; Amendment 1

Date: 22/11/2013

Appendix I: CASP Quality Appraisal Tool

	Critical Appraisal Skills Programme Checklist				
	(CASP, 2002)				
	F = Fully demonstrated (2); P= Partially Demonstrated (1); N= Not demonstrated (0)				
Screening Questions		F	P	N	Comments
		(2)	(1)	(0)	
1	Was there a clear statement of the aims of the research? (Consider: goal, importance, relevance)				
2	Is the methodology appropriate? (Consider: does the research seek to interpret or illuminate the actions and/or subjective experiences of research participants; Is qualitative research the right methodology for addressing the research goal?)				
Detailed Questions					
3	Was the research design appropriate to address the aims of the research? (Consider: has the researcher justified the research design? Have they discussed how they decided which method to use?)				
4	Was the recruitment strategy appropriate to the aims of the research? (Consider: explanation of participant selection, why participants are most appropriate to provide access to the knowledge sought by the study, discussion around recruitment-e.g. why some people chose not to take part)				
5	Were the data collected in a way that addressed the research issue? (Consider: if the setting for data collection was justified; clear how data was collected (e.g. interview/focus groups etc.)?; have they justified methods chosen?; are methods explicit (e.g. exact interview procedure); any modifications of methods? If so have they explained why?; Is the form of data clear (e.g. tapes, videos etc)? Have				

	<i>they discussed saturation of data?</i>				
6	Has the relationship between researcher and participants been adequately considered? <i>(Consider: did researcher critically examine own role, potential bias and influence during (a) Formulation of the research questions (b) Data collection, including sample recruitment and choice of location • How the researcher responded to events during the study and whether they considered the implications of any changes in the research design</i>				
7	Have ethical issues been taken into consideration? <i>• If there are sufficient details of how the research was explained to participants for the reader to assess whether ethical standards were maintained • If the researcher has discussed issues raised by the study (e.g. issues around informed consent or confidentiality or how they have handled the effects of the study on the participants during and after the study) • If approval has been sought from the ethics committee</i>				
8	Was the data analysis sufficiently rigorous? <i>If there is an in-depth description of the analysis process • If thematic analysis is used. If so, is it clear how the categories/themes were derived from the data? • Whether the researcher explains how the data presented were selected from the original sample to demonstrate the analysis process • If sufficient data are presented to support the findings • To what extent contradictory data are taken into account • Whether the researcher critically examined their own role, potential bias and influence during analysis and selection of data for presentation</i>				
9	Is there a clear statement of the findings? <i>• If the findings are explicit • If there is adequate discussion of the evidence both for and against the researchers arguments • If the researcher has discussed the credibility of their findings (e.g. triangulation, respondent validation, more than one analyst) • If the findings are discussed in relation to the original research question</i>				

10	How valuable is the research? • <i>If the researcher discusses the contribution the study makes to existing knowledge practice or understanding e.g. do they consider the findings in relation to current or policy?, or relevant research-based literature?</i> • <i>If they identify new areas where research is necessary</i> • <i>If the researchers have discussed whether or how the findings can be transferred to other populations or considered other ways the research may be used</i>				
----	--	--	--	--	--