Author's response to reviews

Title: Applying an extended theoretical framework for data collection mode to health services research

Authors:

Michael R Robling (roblingmr@cardiff.ac.uk)
David Ingledew (d.k.ingle dew@bangor.ac.uk)
Giles Greene (greene@cardiff.ac.uk)
Adrian Sayers (Adrian.Sayers@bristol.ac.uk)
Chris Shaw (cshaw@ glam.ac.uk)
Lesley Sander (sander@cardiff.ac.uk)
Ia n Russell (i.t.russell@swanse a.ac.uk)
John Williams (j.g.williams@swanse a.ac.uk)
Kerenza Hood (hood k1@cardiff.ac.uk)

Version: 2 Date: 27 April 2010

Author’s response to reviews:

Dear Editor

Thank you for the review of our manuscript, and for the comments of the three reviewers. We are grateful for the detailed and constructive criticism provided and we hope than in preparing this revised version we have improved the quality of the manuscript. Overall we agree with the thrust of the main points across the three reviewers comments and have revised the manuscript in an attempt to address their suggestions and reservations.

Specifically, as you have suggested we have increased the number of examples from health research cited in the text, revised and extended the glossary provided, produced a summary table of key points, and added a more substantial section addressing the increasing role of self-reported outcome measures in routine clinical care. We have also tried to clarify the focus of the manuscript by removing some previous content and better signposting the scope of the article.

Further details of how we have addressed editorial and reviewers comments are provided below. The revised manuscript is provided using tracked changes as requested. This single document includes references, tables and figures. We have addressed nearly all of the reviewers’ comments including major and minor essential revisions and also discretion ary revisions.

We look forward to receiving your feedback on this revised manuscript.

Yours sincerely,

Mike Robling
Response to editor and reviewers’ comments

Editor

Some words, (such as satisficing) are jargon which are unfamiliar to a significant proportion of the health service researchers for whom this discussion paper is presumably intended. A glossary of words and their technical definitions would improve the accessibility of the paper.

• We have revised and expanded the included glossary in table 1.

Consider a table highlighting key points for the different audiences/functions suggested by Reviewer 2 in paragraph 4.

• We have now provided a box summarising key messages (Box 1) at the end of the manuscript

Consider explicitly citing examples from health service research

• We have considerably increased the amount of health related references cited in the manuscript, including more explicit examples drawn from health research literature. This is partially evidenced by the approximately thirty additional references, most of which are set in a health context.

An additional thought is the increasing use of "Patient Reported Outcome Measures" in health care (as opposed in research). You may wish to comment on the relevance (if any) of your discussion in this context.

• We have added a section introducing the use of PROMs in health care within the section 'Particular uses of data'.

Reviewer #1: Catherine Chittleborough

Major Compulsory Revisions

• Thank you for your constructive comments on this manuscript.

2. While the focus of the paper is on how survey mode influences response quality, for it to be useful for researchers in deciding between survey modes, the discussion needs more acknowledgement that a survey mode that results in reduced socially desirable responses, for example, may not be as effective in obtaining high overall response rates. So while you may have respondents that answer truthfully, these respondents may not be representative of your population of interest. Both response rate and response quality need to be
considered in making decisions about data collection modes.

• We have now added a section which addresses coverage and non-response error (in 'Managing node features …').

3. The introduction states that much of the empirical evidence for mode effects has been generated by research outside the health sector. However, I think these issues are well recognised within health research and many empirical studies of differences in various outcomes, from hospital patient satisfaction, to quality of life among various patient groups, to socioeconomic variables within health surveys, exist and thus already highlight that mode is important for health research. This is not the first paper to explore mode issues in health research, and other papers could be acknowledged.

• We have added a substantial number of new references to the manuscript mostly drawn from the health sciences literature. We would certainly agree that there is awareness of such issues amongst some health researchers, although we also believe that much empirical work on this subject is found in other academic / professional domains.

4. The cognitive burden on respondents is one mediator that influences response quality, but what about other burdens? For example, the time it takes to complete the survey, and how completing the survey can fit into other demands on the respondent’s time. Such factors may influence response quality as much as response rates.

• We would suggest that some of these burdens may nevertheless be encompassed by the mediator of cognitive burden. Competing demands may lead to satisficing and we have referred to that issue specifically in that section. We do though recognise that the current model could be further expanded but have tried to locate major elements as they may pertain particularly to health.

5. Are the elements in the social exchange theory, e.g. perception of rewards and costs, and trust in the researcher, influenced by the current climate of increasing market saturation – potential respondents being bombarded with surveys and sales pitches? Again, such issues may have an effect on response quality as well as response rates.

• This sounds credible and we have added a comment to this effect.

6. In the discussion about variables that mediate the impact of mode feature (p9), it is stated that the level of cognitive burden for individuals is less subjective than perceptions of either impersonality or legitimacy, but I am not convinced about this. Is there further evidence to support this statement?

• We have modified the description of this section to clarify that actual cognitive burden is less dependent upon an individual’s psychological appraisal of the survey task, than perceptions of impersonality / legitimacy. We would agree that appraisal still plays a role in the experienced burden – hence the location of this mediator under the general header of appraisal.
Discretionary Revisions

7. The Respondent is described (p10) as the person who is “being assessed”. This sounds as though the respondent is under scrutiny. Perhaps a more appropriate term would be that a respondent is a provider of information? Unless this terminology comes from another source, in which case this should be clarified.

• We have made this suggested change.

8. What is the difference between level of response and validity of response (p18-19)? If the level of an attribute is reported accurately, then the attribute will be valid. Similarly, the accuracy, or validity of a response determines the level of the attribute. While Figure 1 simply lists level and reliability, the text highlights level, reliability and accuracy separately. The text could be re-worded to link the level and accuracy concepts together.

• Yes we agree that this was inconsistent, and have modified both the figure and text accordingly.

9. The concept that year of publication should be considered when comparing survey modes should be introduced before the final summary of the paper.

• This concept is highlighted earlier, in the section on social desirability bias

Minor Essential Revisions

10. The acronym ACASI should be defined when it is first used (p11).

• This definition has been added

11. The sentence “Quality of life is increasingly assessed… etc” (p19, second sentence under Objective/subjective constructs) does not need the comma after “life”.

• Removed

Reviewer #2: Caroline Roberts

Reviewer’s report:

Major Compulsory Revisions

I enjoyed reading this paper. It is a topic I am very familiar with and I believe it is particularly important to bring the literature relating to mode effects in quantitative data collection to the attention of researchers working in domains outside of the social sciences, where factors influencing data quality are perhaps better understood. For this reason, I am keen to recommend this paper for publication and I am advising that the paper be accepted after revision, but I do feel it needs more work than the minor revisions listed below.
The paper was generally very clearly-written and structured and the scope of the literature covered was broad. The model presented provides a concise framework for the reader to begin to grasp the complexity of the problem of mode effects and the various characteristics of data collection modes that can influence the quality of the data collected. I also liked the extension of the model to include additional features not addressed in its original form by Tourangeau et al.

• Thank you for these encouraging comments and for the constructive criticisms provided

Nevertheless, I was left feeling somewhat disappointed by my reading of the paper. Perhaps most importantly, I felt that it did not quite achieve what it set out to, in terms of highlighting the implications of the findings of research on mode effects for researchers working in the health domain. To address this, I think it would be nice to see an expanded version of the section on pages 21 and 22, perhaps incorporating and building on some of the comments raised at the bottom of page 20 about the specific features of research in a clinical context that might render some of what we know about mode effects less applicable. Is there any existing research on this? This seemed to me to be one of the most interesting aspects of the paper, but was never fully developed. If you were to discuss in a more systematic way the specific features of clinical and health-related research that make it different from social surveys, it would allow for a more systematic analysis of the literature - particularly in terms of highlighting where new research is needed. You mention the need for research in various health-related areas, and again I felt this could have been expanded on more - either in the health-related section, or in its own section (e.g. ‘Recommendations for future research’). Without building more on this unique contribution that the paper has the potential to make, I feel that it doesn’t offer much beyond what reviews available elsewhere are able to offer - except, of course, by bringing the literature to the attention of a new audience.

• In our revised manuscript we have aimed to place more focus on the sections specifically addressing data collection in the health sector, and have generally tried to introduce a greater number of relevant references. For example, (and also in line with suggestions with both the editor and with the other reviewer) we have expanded upon data collection in routine clinical care.

• We nevertheless also wanted to emphasise that in many circumstances, issues in collecting self-reported data in a health settings did not really differ from other settings, and have emphasised this point.

• We have re-focused the section on health (now entitled ‘Particular issues in health research’) to emphasise antecedent features, constraints on modes and particular uses of data.

• We have added a new table which summarises how antecedent features (using example of respondent role) may influence response in the context of a clinical trial. This describes potential linkages with mediators such as legitimacy and suggests hypotheses about how these factors may affect response quality.
I was also concerned that the paper started to lose focus as it progressed. By page 23 and 24, where you start to discuss again the characteristics of modes that might influence mode choice, I started to question what the main point of the literature review was. Is it designed to help researchers in the health domain decide how to choose a mode of data collection? Is it about how to assess the quality of data collected in surveys? Is it trying to highlight the problems associated with comparing data from studies that have collected data in different modes? Or is it about the problems involved in using a mix of data collection modes in a single study? These are all important issues, and it is possible that you are attempting to address all of them here. The problem is that it isn’t always clear which one is the primary focus, and so the text becomes confusing to follow. I think it would help to really distinguish between these issues at the start of the paper, and try to make it clearer what the implications of the literature review are for each one (or for those you choose to focus on, if you want to select just a few).

- We accept this criticism, and that we had tried to retain an interest within the paper on too many potential areas. We have changed the Background section to clarify what the main focus of the manuscript is, and importantly what we are not attempting to address (for example it is not an attempt to review the use of mixed mode approaches in surveys).

This point about clarifying the focus is also important because at times new issues crop up but are never fully expanded on. For example, you mention possible advantages of switching modes in a single study on page 12, but you never really discuss or highlight the problems involved in mixing modes, though these are implied throughout. Similarly, you raise the problem of mode effects on survey participation (the fact that modes influence people decisions about whether to participate), but this is quite a separate issue from the effect on response quality (except where the mode acts as a common cause of both nonresponse errors and measurement errors, which you mention at the top of page 13). It would be good to elaborate more on some of these points, because at present, I’m concerned that you run the risk of oversimplifying some important issues.

- Similar to the point above, we accept this criticism and have revised the manuscript to make it clear where we will expand on issues, and where we feel it is not appropriate to do so (because of the inadequacy of space within the manuscript / potential to divert the reader’s attention). We have also removed some of the original text which we felt did not contribute to the aims of the manuscript.

- Minor Essential Revisions
Page 5, 2nd para, lines 3 and 4: give references here

- We have added appropriate references here

Page 8, 2nd para: Is this correct? My understanding is that Tourangeau et al.
distinguish 3 features rather than 4 - there seems to be some overlap here.

- There is some inconsistency across the two references originally cited, so we have now retain only a single (the more recent) reference which does refer to four features.

Page 9, end of 2nd para: include a reference for satisficing

- A reference has now been added here

Page 9, 2nd line from the bottom: there’s a grammatical problem here

- We have modified this sentence

Page 10, end of para 1: check ‘an exhaustive listings’

- We have modified this sentence

Page 14, end of para 2: do you have any references relating to ‘pace’ that you can insert here?

- We have added a citation to Tourangeau at this point, and also to Krosnick.

Page 15, top para: similarly, a reference relating to multitasking here?

- We have added a reference to Krosnick for this point

Page 15, 2nd para, 5 lines from bottom: should this be ‘pauses’?

- Yes, now modified

Page 15, 2nd para, last sentence: do you mean the satisficing model? do you mean specifically in the health domain? There are several studies that have applied it in the social sciences (including some by myself and colleagues!).

- We have added a reference and illustrative example (from social survey domain). This was not intended to be restricted to health at this point.

Page 19, top of page: it’s not really clear here whether you’re talking about reliability across measures in a survey or across different surveys. It seems all three mode features could influence both reliability and validity.

- At this point we were directly drawing upon Tourangeau’s model – and we have now made this clear in the revised text.

Discretionary Revisions

Page 6, last sentence before Discussion: is the ‘systematic method variance’ referred to here different from biases such as social desirability and acquiescence bias?

- We have modified the presentation of this section generally now and excludes the use of this descriptor
Page 6, third to last line: ‘greater efficiency’ - compared to what?
• The comparator is now described

Page 7, 1st para: you describe differences between face-to-face and telephone questionnaires, but it doesn’t seem clear that the differences arise as result of the mode characteristics. The last sentence also seems a little vague - what exactly do you mean by ‘how the survey is appraised’? Can you expand on this? It’s not clear what is implied.
• We have modified this description to better clarify

Page 7, 2nd para, 1st line: clarify what the ‘original distinction’ is.
• We have now changed the initial sentence

Page 8, subheading: it’s not really clear at this point whether you are talking about the effect of mode on participation in surveys or on response quality.
• Sub-heading modified

Page 9, 2nd para: there is considerable evidence that different characteristics of the survey can influence legitimacy so I wonder whether it is more objective than you imply. Equally, there are individual differences influencing burden too - e.g. intelligence, topic-related knowledge, interest/involvement in the survey topic, etc.
• We have re-worded this section. What is subjective is how individuals appraise stimuli (even if their response is general consistent) – ie it is a subjective process.

Page 9, top of para 3: which response quality indicators?
• We have added the indicators

Page 11, last para: again I feel it’s worth mentioning that there are different ways of establishing legitimacy in surveys but that the mode may influence how it is perceived.
• We have not directly addressed this here, but have suggested in table 2 ways in which legitimacy may be affected in a trial setting

Page 12, end of para 2: How do leverage and salience interact with mode and what is the outcome? I feel it isn’t quite clear why this is important given the discussion up to this point focuses mainly on the effect of mode on response quality.
• We have added some further description at this point

Page 12, half-way through last para: it’s not clear what ‘switching between modes’ means here - do you mean to follow up nonrespondents as in a sequential mixed mode design? This could equally have the effect of irritating
sample members!

• We have amended the description at this point

Page 17, end of para 2: does this mean the environment at the time of questionnaire completion?

• We have added to the description to clarify

Page 17, section on acquiescence: I’m not entirely convinced by the decision to treat acquiescence differently from other satisficing effects, although I take the point that it shares similarities with social desirability bias. Can you make a stronger argument here?

• We have not expanded on this point further, for brevity sake

Page 19, 1st para, last sentence: it’s not totally clear what this means.

• We have omitted this description in the revised text

Page 20, top of page: the important point here seems to be that ‘objective’ behavioural measures are in principle, objectively verifiable.

• We agree

Page 20, para on respondent role: is it true that this research has mainly been on student samples? I thought there was quite a lot of research from population surveys?

• We have clarified this point and made it explicit that this draws upon relevant meta-analyses where many included studies did involve undergraduate students

Page 20/21: section on respondent characteristics: this seems a bit speculative in places. As mentioned, legitimacy depends on ‘objective’ things like survey topic, sponsor, etc., rather than respondent perceptions. Similarly, the likelihood of satisficing seems to depend partly of the design of the questionnaire (and as you’ve mentioned, the mode), so I’m not sure you can speculate about differences between different survey populations.

• We have included additional supportive evidence for some of these points (e.g. that perceived personal gain and benefits may be influential).

Page 21, 2nd para: self-reports of sexual activity also depends on respondent sex (men overreport, women underreport).

• We would agree but just have listed these characteristics as an example, and not intended as an exhaustive list

Page 22: last para: questionnaire construction only really becomes important when you are combining modes in a single study, but this hasn’t really been the focus of this paper.
• We have removed this section form the manuscript.

Reviewer: Andrew M Garratt

Reviewer’s report:

Major compulsory revisions
None

Minor essential revisions

• Overall, thank you for these constructive comments

Abstract summary

Health related research is quite different from market research where much of the research has been conducted relating to mode features and response quality. Therefore "...is of importance to health research" is perhaps better than "equal importance". The same applies to the Summary where "generally applicable" is perhaps more accurate given the current state of knowledge than "equally applicable".

• Both wording changes made

Background

Starting the first sentence with 'subjective outcomes' is perhaps being too specific given the general nature of the article. Health related surveys of patient, practitioners and the public can include a range of issues from health behaviours and lifestyle through to health care and illness experiences, satisfaction and outcomes. An overview of the main types of health survey (study objectives/measurement and survey participants) might give the article greater context as well as being useful for considering the potential importance of the different mode effects in relation to the different survey types.

• A new initial paragraph has been added to better set into general context the focus of the review.

Background para 2. A reference is needed to support the statement that there is considerable evidence in relation to survey methods and response rates.

• Supportive references are now cited

Page 13 'Taking the easy way out' I am not sure what is meant here by 'satisfactory answer'. The sentence could simply end with "...in giving a response".

• Suggested wording change made

Page 19, para 1. I would also advise against using 'outcome measure' as it is not such an obvious example as health behaviours and lifestyle.
Page 19. Measurement construct. I am uneasy about the statement that QoL is increasingly assessed as an individually-defined subjective construct with an obscure reference from 1996. The vast majority of what are referred to as quality of life measures are standardised disease-specific measures and standardised generic measures of health such as the SF-36 and EQ-5D. 'Instrumentation bias' is introduced here without an explanation of what it is.

• We have modified this description as part of a general revision of this section

Page 20. Respondent characteristics. Final para. An example from the measurement of patient experiences and satisfaction might be useful here. It is fairly well documented that patients completing questionnaires measuring satisfaction with care have higher rates of satisfaction if they complete the questionnaire at the clinic compared to at home (Crow et al, HTA 2002).

• We have added this example to the text

Page 21. Mode-feature effects in health Given the importance of health outcomes measurement as a form of patient survey it might be worth mentioning Computer Adaptive Testing. CAT is arguably the most advanced method of administration available resulting in lower response burden and the tailoring of questions to the individual patient’s health state. It can also be applied to other types of health survey.

• We have now added a brief section introducing this concept / approach

P25 Summary. The Abstract has two recommendations relating to greater consideration of and clarification of mode features which are not included here. Greater consideration of how features of different methods of data collection affect responses is necessary but we also need further research into potentially important mode features of health surveys. Piloting of alternative mode features might be recommended when there is uncertainty relating to survey design.

• The newly added summary box (and table 2) provide additional specification of the need to consider & clarify mode feature effects.

Discretionary revisions
P 21 para 2. I suggest replacing 'emphasise' with 'have'.

• Suggested wording change made