Comments from the authors

Comments of the Associate Editor:
In particular, the revised version should focus to:
- provide theoretical background re the hypotheses
- make more clear how analyses are done (see review of referee A)
- provide some more information about the questionnaire (reviewer A)
- remove details about data handling (reviewer A)
- take the remarks of reviewer B about costs into account
Furthermore, limit the length of the paper to maximally 10 pages instead of 14, including references.

Authors' comments: We have complied with these issues in the revised manuscript which will emerge from our comments to the review comments. We thank the editor and the reviewers for their constructive suggestions to the manuscript.

Reviewer A
The paper compares postal and web-surveys with respect to response rates and data quality, as indicated by the number of don't know answers and the level of item nonresponse. In my opinion, the effects found and their interpretations seem to comprise only a modest contribution to the survey research methodology literature. The paper lacks a coherent theoretical overview and review of the literature. The hypotheses are not clearly explained with arguments and evidence based on earlier findings and theory. Thus, the thinking in the paper is less carefully and thorough than it might be so that it is really not clear what the results mean.

Authors' comments: More theoretical and methodological background for the study is provided in the revised manuscript.

Although here and there some hypotheses are presented, they are not well-argued. The authors just explain what they did ("two versions were compared to test whether the Web-based survey had a lower response rate than the postal survey"), but they do not explain why it is important to test this, and why this was expected in the first place. In general, not many arguments and expectations are given in the paper. In some cases references are clearly lacking. For example, the statement "web-surveys have also been claimed to provide lower respondent errors and to increase the completeness of response". The authors do not state who is claiming this, and do not explain why this could be expected; what is the theoretical rationale for believing web-surveys lowering respondent errors? In addition, to me it is not clear what "increasing the completeness of response" means; does that refer to lower item non-response, more complete answers to open-ended questions or perhaps something else?

Authors' comments: More clarification and background is provided as well as hypothesis.

The authors also state (citing Cook et al) that 'representativeness of the sample is much more important than the response rate'. Although I can see their point in this statement, I think the authors should be aware that the three variables that they use (age, gender and region) are not enough to judge how representative a sample is. Even if more variables had been available, you never know whether the samples are not representative with respect to a relevant variable that was not measured. Therefore, a finding that the item non-response was lower in the web survey than in the postal version, may still be due to differences in the non-response rates. Although gender, age and region were controlled for, there may be an unobserved variable explaining why respondents filling out the web survey were less likely

to skip an item than respondents filling out the postal survey. This makes it difficult to conclude whether it is a mode effect (i.e., web surveys are easier to fill out) or a selection effect (more motivated respondents in the web survey than in the postal version).

Authors' comments: We have added these relevant comments to the paper.

It is not clear how the analysis was done. A two sample proportion test was used, in which two binomial proportions can be compared. It is not clear which two percentages were compared, and how this was done for the non-binomial distributions (age and region).

Authors' comments: The calculation has been clarified.

The main finding, with respect to differential response rates, is the lower response rate in the web survey than in the postal survey. Firstly, this is not a new finding, as the authors also note, there is a growing literature on web survey methods. However, they do not review earlier findings with respect to response rates in comparisons of web and mail versions of the same questionnaire, nor do they relate their findings specifically to earlier findings.

Authors' comments: We have added references to relate our result to the literature.

What is relatively unique in this study, is that all respondents were contacted in the same way, by regular mail, and that information was available on all individuals contacted. As state Esteban Mccabe et al (2002) state "In most cases electronic mail (email) is used to communicate with respondents, and the Web is used for the actual data collection.". See also Denscombe 2006. While this appears to be an attractive aspect of the study, it also entails a danger with respect to the interpretation of the research findings. The authors explain lower response rates for web surveys only very briefly, and are only focusing on internet access and familiarity with computers as explanatory factors. However, they do not address the point that in this particular study, the very fact that in the web survey a different mode was used to contact respondents, than the mode used for data collection may explain the reduced response rate. When respondents want to comply with the request to participate, they have to switch modes; i.e. they have to take the letter to a computer, and type in the Web link, before they can actually fill in the questionnaire. This is of course different from the common request by email, in which respondents do not have to switch modes, since often can follow a link or copy and past the link in the web browser. Thus, the additional effort of the mode switch and the chance of respondents giving up because they made a typing error, may explain a lower response rate in the web survey.

Authors' comments: We have added these relevant comments to the paper.

The paper could also contain more details about the questionnaire, for example including (part of) the actual question wordings in an appendix would have been very helpful. That may clarify what the authors mean with their note that 35 closed-ended questions were used, some with multiple answering categories. In contrast, some information could be omitted (e.g. about data handling such as the electronic scanning of the postal questionnaires and data extraction from the host web server), as it is not very relevant to discuss such details.

Authors' comments: We have added a bit more information about the questionnaire and deleted some of the information about the data handling process.

Finally, the paper is not very well written; it shows some awkward sentence structures like "number of nursing homes increasing by 20,000 up till 2020. " or "the questionnaire aimed to explore the future potential residents about their attitudes".

Authors' comments: We have had a native English speaking person to go through the revised manuscript.

References

Balter, K. A., Balter, O., Fondell, E.& Lagerros, Y. T. (2005) Web-based and Mailed Questionnaires: A Comparison of Response Rates and Compliance. Epidemiology. 16(4):577-579, July 2005.

Bosnjak, M. & Tuten, T.L. (2001) Classifying Response Behaviors in Web-based Surveys. Journal of Computer-Mediated Communication, April 2001

Denscombe, M. (2006)Web-Based Questionnaires and the Mode Effect: An Evaluation Based on Completion Rates and Data Contents of Near-Identical Questionnaires Delivered in Different Modes Social Science Computer Review 2006; 24; 246 DOI: 10.1177/0894439305284522

Esteban Mccabe, S., Boyd C., Couper, M. Crawford, S. & D'Arcy, H. (2002) Mode Effects for Collecting Alcohol and Other Drug Use Data: Web and U.S. Mail Journal of Studies on Alcohol , Vol. 63, 2002

Authors' comments: Thank you for the good references

Reviewer B

Summary:

The paper deals with a large size mail survey of persons aged 50-75, where the persons – by random allocation – either received a traditional postal questionnaire or a letter with a Web link to an online version of the same questionnaire. The two surveys are compared with respect to 1. response rate, 2. non-response selectivity with respect to age, gender and region – available from the sample frame, 3. the distributions of income, education and health status of respondents, 4. item non-response and 'don't know' answers and 5. self perceived difficulty of DCE questions.

The authors conclude that the Web based version had lower response rate, mixed results in selectivity and higher cost per response but has better data quality (in terms of item non-response and don't knows).

General Comments:

In my opinion the paper is very clear. It is well organized and well written. At the same time the contribution of the paper is rather limited. The paper lacks theory. The results refer to the situation in Denmark, to older respondents and – most importantly – to only a small set of 'basic variables' (age, gender, region, income, education) and their relationship to non-response. The paper appears to me as a research note. As such I would find the results worth publishing.

Detailed comments:

Table 3: The binning of Income seemed initially a unfortunate due to small lower and upper income classes, but shows very well how poor the web based survey does in the lower income class.

Authors' comments: This conclusion has added to the manuscript

Table 4. Since it is not necessary to restrict the analysis here to the variables that are from the sampling frame, I would be interested in adding education, income, health and civil status as predictors

Authors' comments: We have restricted the analysis to the variables which we had information on from the initial draw from CNR since the variables from the survey can be missing which will cause a catch 22 situation – the confounding factors may themselves have item non-responses. We have added this remark to the manuscript.

Table 5. It is not clear to me why the topic of perceived difficulty of DCE questions is in this paper. Do you see this as an indicator of data quality?

Authors' comments: We have decided to delete this table and the analysis of the perceived difficulty of the DCE questions since it is not a core issue in the paper, and because this variable cannot be interpreted as a general indicator of data quality.

Discussion I found the evaluation in terms of costs very interesting. You clearly show that in your design the benefits of using a Web survey is not in terms of costs. How would the costs have turned out in case of equal response rates? I suggest that you evaluate the cost of a response of Web relative to the cost of a response in the postal survey (as ratios), since the use of so many currencies (US$, €, DKK and £) is confusing. In presenting the results, take the postal survey as reference and discuss to what extent the web survey does better or worse. The discussion of Table 3 on page 7 is the other wayaround.

Authors' comments: The idea of presenting the cost of survey modes as ratios is very good and is adopted in the manuscript. We have furthermore calculated the unit cost per reply in case of equal response rates.

Title: 'Differential response rates in postal and Web-bases surveys in older respondents'.
I suggest 'among older respondents'.

Authors' comments: We have adopted this suggestion