Author’s response to reviews

Title: Is web interviewing a good alternative to telephone interviewing? Findings from the International Tobacco Control (ITC) Netherlands Survey

Authors:

Gera E Nagelhout (gnagelhout@stivoro.nl)
Marc C Willemsen (mc.willemsen@stivoro.nl)
Mary E Thompson (methomps@uwaterloo.ca)
Geoffrey T Fong (gfong@uwaterloo.ca)
Bas van den Putte (S.J.H.M.vandenputte@uva.nl)
Hein de Vries (Hein.deVries@gvo.unimaas.nl)

Version: 2 Date: 26 May 2010

Author’s response to reviews: see over
**Author’s response to reviews**

**Title:** Is web interviewing a good alternative to telephone interviewing? Findings from the International Tobacco Control (ITC) Netherlands Survey

**Authors:**
Gera E. Nagelhout (g.nagelhout@stivoro.nl)
Marc C. Willemsen (mc.willemsen@stivoro.nl)
Mary E. Thompson (methomps@uwaterloo.ca)
Geoffrey T. Fong (gfong@uwaterloo.ca)
Bas van den Putte (S.J.H.M.vandenputte@uva.nl)
Hein de Vries (Hein.deVries@GVO.unimaas.nl)

**Version:** 2
**Date:** 29 April 2010

Thank you for asking us to revise our paper. We have addressed the comments made by the reviewers, and made improvements to the English language. Our responses to the individual reviewer comments are listed below.

**Reviewer 1: Michael Erhart**
We have addressed the comments made by reviewer 1 and marked them yellow in the paper.

**Major Compulsory Revisions**

*Page 4, para 1:* It would be important to also contrast web-interviewing with paper and pencil administration of questionnaires.
Authors: We have added that in comparison with paper-and-pencil interviewing, web interviewing requires no data entry, presents one interview question at a time, and can use complicated skip patterns.

*Page 4, para 2:* It does not come totally clear what the sampling error refers to. It also needs to be explained what the term probability sampling refers to and why self-selection precludes estimation of the sampling error.
Authors: On page 4, we have added a more elaborate explanation of sampling error: due to sampling error, point estimates of a sample study will not be precisely the same when another sample or the entire population is surveyed. Also, we have explained why self-selection leads to unreliable confidence intervals: Self selection may result in a systematic bias, and the associated variability cannot reliably be described with a confidence interval.

*Page 4, para 2 to Page 5, para 1:* This presentation of 'error sources' could be supplemented by introducing a generally important dimension along which error could be classified: Precision versus Bias.
Authors: Sampling error and measurement error are usually thought of as having both bias and variability components. We have not introduced this classification of error formally, but we allude to both bias and variability in the discussion of sampling error and measurement error.

*Page 5, para 1:* This passage could be supplemented by some statements on differential item / test functioning across modes of administration: This would not only concern the question about different measurement (on average) results but also e.g. if the measure (and their indicators/items) are functioning in a similar discriminative, reliable and valid way across the different modes of administration.
Authors: We have added that there are indications that telephone interviewing results in less reliable and valid measures than paper-and-pencil interviewing and in more complicated factor structures than web interviewing.

Page 7, para 1, fourth last line: Please also provide the response rate as defined in page 8, para 1, last three lines. However this information would be better placed in the results section.
Authors: We have replaced the information about cooperation and response rates to the results section and explained why we could not give a response rate for the web survey.

Page 7, last para, lines 1-2: It would be important to know how the TNS NIPObase was recruited, which are the inclusion and exclusion criteria, and why it is only representative for region.
Authors: More information about TNS NIPObase and its recruitment procedure is added to the methods section.

Page 7, last two lines: The authors wrote “In this way, every member of the Dutch population with a fixed line telephone had the same chance of being sampled for this survey.” This is not totally true. Only every member of the TNS NIPObase had the same chance of being sampled.
Authors: Every member of the Dutch population with a fixed line telephone had a chance of being sampled for this survey proportional to the number of telephone numbers in TNS NIPObase which “agreed” with the population member’s number in all but the last two digits. Since this is not the same as every member of the Dutch population with a fixed line telephone, we have deleted this from our paper.

Page 8, para 1, last three lines: It remains unclear what the “response rate” and “cooperation rate” refers to. Please indicate this in a half sentence / sentence. These information refer to results of the study and thus would be better placed in the results section.
Authors: We have explained more about the response rate and cooperation rate in the results section.

Page 8, last para: This is a creative approach to test coverage error. Yet still the confounding between both coverage errors (web and tel. mode of administration) cannot be totally controlled.
Authors: We could have studied the different error sources better if we had randomised respondents into the web and telephone survey. However, our approach has more ecological validity, since we used a real study with a commonly used design.

Page 8, last para, line 1-6: These sentences could be moved to the statistical analysis section.
Page 9, para 1, line 1-8: sentences could be also moved to the statistical analysis section.
Page 9, para 3, line 1-4: These sentences could be also moved to the statistical analysis section.
Page 10, para 2: This section could be enhanced by including the information mentioned in the comments above. It would be helpful to make more salient what the different analyses aim at.
Authors: We moved these sentences to the statistical analysis section and restructured the statistical analysis section.

Page 10, para 2 and before: A fundamental Problem concerns the fact that in its current from the statistical analysis strategy does not permit to examine measurement error independent from coverage and response error, as the latter were not controlled. However if there are any differences in the socio-demographics or socio-economic make-up of the web and tel. Based samples, it is likely that these differences would also manifest in differences in the measurement results between the different modes of administration. One possible solution would be to include such demographic and socio-economic variables as covariates in the statistical analyses. This however would demand applying e.g. multiple regression models. In it s current from the manuscript on several times examine the question: Are there any differences between the samples – however it is not possible to identify which proportion of difference/error is attributable to which source of error between the samples. This is a major shortcoming of the work.
Authors: We thank the reviewer for his statistical advice. We have added multiple linear regression analyses to study measurement error when controlling for demographic variables.
Another major shortcoming refers to the rather sparse statistical analyses: In fact the authors examined only omnibus differences in the outcome distribution. In the ordinal scaled variables there were no attempts to examine for differences in the central tendency (e.g. U-Test); larger floor or ceiling effects (e.g. logistic regression) etc. In the interval scaled variables no attempts were made to examine for differences in the dispersion of the answers (e.g. Leven tests). It would be also interesting to see if there are any effects of the mode of administration on the strength of association between certain variables (e.g. gender differences in smoking – are they more pronounced under a particular mode of administration?). Or are the items answered in an internally more consistent manner under a particular mode of administration? The internal consistency or the association between the different smoking questions could be computed and compared across samples (e.g. Feld Test for differences in Cronbach alpha). Interaction between mode of administration and one outcome with regards to the prediction of another outcome.

Authors: We have added Mann-Whitney tests for the ordinal variables and Levene’s tests for the interval variables. Also, we studied interactions of mode of interviewing and educational level with multiple linear regression analyses, to see whether the known association of smoking cessation determinants with educational level was present in both interviewing modes. We believe that these shortcomings of the study are now dealt with.

Discussion: Of course the discussion is generally limited by the shortcomings mentioned beforehand. Page 13, last para: See comments above: there is no valid conclusion possible on the measurement error without controlling for coverage and response error.

Authors: The multiple linear regression analyses resulted in the same conclusions as the univariate analyses. Therefore, our discussion of the results did not change much. We have deleted the sentence in which we explained that the differences in attitude towards smoking and self efficacy for quitting could also be caused by the higher educational level of telephone respondents. We have added a discussion about the observed interaction of educational level with interviewing mode on self efficacy for quitting.

Page 14, first lines: This statement taps an important issue. I wonder if the authors tried to examine the effects of gender and age differences in the interviewing person on the outcomes.

Authors: Unfortunately, we did not have information about the interviewing persons.

Quality of written English: Needs some language corrections before being published

Authors: We have made improvements to the English language.

Reviewer 2: Valentina Rosata
We have addressed the comments made by reviewer 2 and marked them blue in the paper.

Minor Essential Revisions
The main message of the present manuscript is that web interviewing can be a good alternative to telephone interviewing, since there are indications that “web surveys may even obtain better data quality than telephone surveys”. Authors should add to the Discussion section at least a paragraph suggesting how to reduce the limitations of the web surveys, including the relatively high number of "Don’t know" in web interviewing.

Authors: We have added a paragraph in the discussion in which we discussed this issue.

Authors could discuss strengths and limitations of a mixed-mode approach.

Authors: We have also added a paragraph in the discussion in which we discussed this issue.

Discretionary Revisions
Please, expand the acronym ITC at the first occurrence in the Abstract.

Authors: We have expanded the ITC Acronym.
Results section of the Abstract: please, clarify that the first sentence refers to the web sample and the second to the telephone one. Results should provide some data, including the cooperation rate according to both samples, and the main findings.
Authors: We have added this to the result section of the abstract.

In the methods, the authors mentioned the amount of reimbursement for each survey. How did respondents receive the compensation? Please specify.
Authors: We have added to the methods section that web respondents earned points for every answered question which they could exchange for money, as is standard procedure in the TNS NIPObase web panel. Telephone respondents received their reimbursement by mail after completing the survey.

Do you have information on type of internet connection for your web sample? I believe that the generalization of your findings to other countries may be influenced by differences on availability of high-speed line. Please further consider this in the Discussion section.
Authors: Unfortunately we do not have this information of our sample. However, 88% of Dutch internet users have a broadband connection. We have added this to the limitations section.

The #2 and t values are uninformative, I suggest to remove them from tables, leaving the p values, only.
Authors: We have added the p values in the text of the results section.

Authors could consider Clark et al., 2008, PMID: 18312649.
Authors: Clark et al. (2008) studied mode differences between paper-and-pencil surveys, telephone surveys, and computer assisted surveys. They found few meaningful differences that were not consistent. We have added this reference to the introduction section.

Response rates (named “cooperation rates” in table 1) for both surveys are relatively high. This deserves to be considered in the Discussion section. However, comparisons between the two samples are biased by the different reimbursement given in the two surveys.
Authors: We have added to the discussion section that cooperation rates were high for both the web and telephone survey, which can be explained by the use of a well respected market research company and the use of reimbursements.