Reviewer's report

Title: Effects of phone versus mail survey methods on the measurement of psychosocial constructs in adolescents

Version: 1 Date: 14 May 2009

Reviewer: Herbert Fliege

Reviewer's report:

The aim of the study was to investigate how self-ratings and proxy ratings of adolescents’ health-related quality of life and some other psychological and behavioural aspects differed between different modes of questionnaire administration, i.e. either by telephone or by post. The research question is specifically important in the area of quality of life research where large surveys need to be carried out and where phone interviews offer a cheaper and quicker means of raising data and yielding immediate data sets with fewer missing values than traditional mail surveys.

The study made use of a representative sampling design. The two selected questionnaires are representative of an established standard of QoL instruments often used in large cross-national surveys in adolescents. The statistical analyses are adequate in order to answer the research questions.

The main finding is that there are significant and meaningful differences between mean QoL scores yielded by the two different methods. On the whole, phone surveys resulted in more positive QoL mean values than mail surveys. At the same time, and perhaps more importantly, there were indications of a lower reliability of the data gained from phone surveys.

The report is clear – though the text would gain from some native English sub-editing – and the presentation of the data is adequate. The discussion adds to the interpretation of the findings and makes some suggestions how the found differences should be dealt with.

I have two more important suggestions and a couple of minor issues.

Major suggestions:

1. Methodology effects for phone surveys as compared with mail surveys are highly relevant. In my opinion it is important to investigate these effects in more detail. Concerning the study at hand I therefore suggest adding an analysis of differential item functioning (DIF). Do items perform differently – with respect to measuring the same construct of the scale – depending on whether the item was administered by phone or post? By adding a DIF analysis the study could attain better insight on the question if the phone/post effect is a general effect occurring in any QoL measurement or if it is due to specific items or even the way items are worded.
2. I think that the advice offered in the discussion on how to amend the methodology effect (to subtract the empirically determined differences between means from future scores assessed via phone) seems well-meant but somehow weak as long as it is not tested what it would actually attain. This can only be carried out in a separate study. Furthermore this would not do away with the problem of a diminished reliability. This may be mentioned in the discussion/conclusion.

Minor issues:

1. Title: What are “psychosocial constructs”? To my mind this is rather a vague term, suggesting everything but actually saying little. Could the authors be persuaded to change this to a more specific term like “health-related quality of life” or “subjective health” or the like? (this applies also to first paragraph of the abstract)

2. P. 7 There was a difference in sampling availability between the phone and the mail sample that results from phone number availability. Could phone number availability in any way be correlated with the outcome variables? E.g. could families with a lower QoL be less prone to provide their phone number? Phone calls may be considered more intimate. Furthermore, since participants were first addressed via mail anyway there was no need for them to provide their address, whereas providing the phone number required an additional deliberate act. Could this lead to a selection criterion that is associated with QoL? If so, it should be mentioned in the discussion as it may apply to other phone studies, too.

3. P. 12 reports that differences in KINDL scores between mail and phone surveys were slightly pronounced for the proxy version than for the self-report version. Was there a significant interaction effect? If so it could be included in the text.

4. P. 14 “phone yielded higher problem rating” # I suppose this should be “lower”. Starting the next sentence but one with “however” does not make sense to me. I do not see that there is a contradiction. Does it not point to the same direction?

Level of interest: An article of importance in its field

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

Statistical review: No, the manuscript does not need to be seen by a statistician.

Declaration of competing interests:

I declare that I have no competing interests.