Author's response to reviews

Title: Handheld Computers for Self-Administered Sensitive Data Collection: A Comparative Study in Peru

Authors:

Antonio Bernabe-Ortiz (03887@upch.edu.pe)
Walter H Curioso (wcurioso@u.washington.edu)
Marco A Gonzales (marcgonzales@hotmail.com)
Wilfredo Evangelista (willyuni@yahoo.com)
Jesus M Castagnetto (jesus.castagnetto@gmail.com)
Cesar P Carcamo (carcamo@u.washington.edu)
James P Hughes (jphughes@u.washington.edu)
King K Holmes (worthy@washington.edu)
Patricia J Garcia (pattyg@u.washington.edu)

Version: 2 Date: 22 November 2007

Author's response to reviews: see over
Dear editors,
Dear reviewers,
BMC Medical Informatics and Decision Making

I’m writing, on behalf of my co-authors Walter H Curioso MPH, Marco A Gonzales, Wilfredo Evangelista, Jesus M Castagnetto PhD, Cesar P Carcamo PhD, James Hughes PhD, King K Holmes PhD, Patricia J Garcia MPH, and myself to response the comments of reviewers about the manuscript ID 1032255907152872

Responses to the First Reviewer:

General:
The title was changed to: “Handheld computers for self-administered sensitive data collection: A comparative study in Peru”

Major compulsory revisions:
Page 3, line 7: “Reporting of sexual behaviors…”
This change has been made. Now, the sentence is: “Social desirability or self-presentation interviewer can affect reports about sexual behaviors as well as other sensitive behaviors. This might change the analysis for non-responses items”.

Page 5, paragraph 1 under header: “The handheld computer program”
This change has been made. Now the header is: “Program used in handheld computers”.

Page 8, line 11: “This agreement was higher in those…”
This change has been made. Now, the sentence is: “The agreement in those who had completed at least high school was higher than those who did not”.

Methods:
It appears from the discussion that drop-list/lookup lists were used on the PDA but this is not mentioned in the methods.
This part was added in the methods section of the study: “The questionnaire contained a set of data entry types (pop-up lists, multi-option answers, etc). Participants entered data using those types of entry options. They chose answers from a list previously established. Participants did not have to entry text using the pen stylus”.

Unclear why the PDA was concealed (in the Methods) although mention is made in the discussion
We agree with the reviewer. But, we do not believe that the fact of concealing the PDA with Styrofoam shield affect to the results of the study. PDAs were concealed due to the risk of being stolen during the fieldwork; as a result, the reason is only mentioned in the Discussion section.

Page 6, paragraph 2: are these statistical tests appropriate given that the data compared are by the same observer, on the same data?
The analysis was performed by a Peruvian statistician and then reviewed by an American statistician. Both of them agreed the analyses were correct. Similarly, previous studies performed the same kind of analyses for similar surveys.

How were participants divided for the second survey? Was this a random process?
Field workers were assigned to teams of two alphabetically based on their last name. Within each team, the first interviewer conducted the interview with the electronic format while the second interviewer conducted the interview with the paper format. As a result, half of participants answered the PDA questionnaire and the other half responded the paper-based questionnaire. This part has been added in the methods section.

**Discussion:**
The discussion is lengthy and could be shortened.
Discussion was has been shortened.

The “three important strengths” (last paragraph, page 9) are largely tautologic and could be deleted.
This paragraph has been shortened.

Page 8, 1st sentence, 2nd paragraph: what is meant by “rates of response”? In this case rate of response was referred to the fact that the same proportion of participants answered “yes” or “not” independently of the method used. Now, the sentence is: “In the second survey, data collected by both techniques were very similar, which is supported…”

Page 8, last paragraph “When comparisons were performed…” should probably mention that the rates were lower in the PDA group.
The change has been made. Now, the sentence is: “When comparisons were performed to evaluate data accuracy through the number of missing values and inconsistent answers, these were statistically lower in the PDA group”.

Page 9, line 2: “6.9% of interview” is not mentioned in the results.
This result was added in the results.

**Responses to the Second Reviewer:**
**Major compulsory revisions**
What pool of individuals was used to find potential participants? What selection criteria were used? What exclusion criteria, if any, were used? How many potentials participants were approached, but decline to participate?
The next paragraph has been added in the study design and setting section to clarify these questions: “In both surveys, a sample of clusters was selected; then a census of each household in the selected clusters was conducted. Within each household, eligible individuals (male or female, 18-29 years, literate, and in the household at the moment of the interview) was selected. Participants provided verbal informed consent prior to participate, and completed a detailed questionnaire on sexual practices. Participation in both surveys was anonymous”.

The first study is biased in that all participants were asked to complete the paper-based survey first. We recognize this bias as a limitation in the discussion. The following sentence has been added in that section: “Some bias could have been introduced in the first survey because all the participants were asked to complete the paper-based before PDA questionnaire. However, we believe that whether the half of participants had firstly responded to the PDA questionnaire, they would not have paid attention to the paper questionnaire or would have left it without answers due to the boredom caused by answering the questions twice, which would have been more unfavorable to the paper questionnaire”.
The second survey divided participants into two groups by “convenience”. It is not clear what this means, and the rationale for not randomizing the participant into the two groups should be clarified. Field workers were assigned to teams of two alphabetically based on their name. Within each team, the first interviewer conducted the interview with the electronic format while the second interviewer conducted the interview with the paper format. As a result, the half of participants answered the PDA questionnaire and the other half responded the paper-based questionnaire. The term “by convenience” was changed in the paper for the explanation above mentioned.

Which subset of questions did the authors use for analysis, how did they select this subset, and could this subset have contributed to any bias in the results?
We selected the most sensitive questions about sexual behavior. Demographic, labor, residence mobility questions, and knowledge about STDs included in the questionnaire were not used for this analysis because they were not considered sensitive. We do not believe that the questions selected for the study contributed to any bias in the results.

The major flaw in the study design is that if recall bias. But it should be explained more clearly than the authors did.
We agree with the reviewer. The recall bias is an important limitation in this study. But also, it is a common bias in this kind of studies. Unfortunately, this cannot be changed because the study has been completed. However, some changes have been made in the discussion to define how this could affect the results.

It would be helpful if other potential confounding factors and issues were discussed.
We also evaluated educational level as a possible confounder whose results are shown in the paper, but no difference was found. After that, we do not believe that there were other potential confounding factors.

What effect, if any, did the training session have on the participants? Do the authors feel that a training session is required (or practical) if this approach were used to collect data or other studies?
Since young Peruvian people are not familiar with the use of handheld computers, rather than desktops and Internet, we decided to conduct a short training session before collecting data. In addition, we conducted the training to recognize the type of possible models of questions and responses, and to avoid PDA screen damage by pressure. A comment about this has been added in the Discussion section.

Do the authors feel the Styrofoam shield (which seems to be usual), would be practical for others and create confounding factors for their study?
We do not believe that the fact of concealing the PDA with Styrofoam shield affect to the results of the study. PDAs were concealed due to the risk of being stolen during the fieldwork; as a result, the reason is only mentioned in the discussion.

Why did the authors use the Bonferroni’s procedure to adjust those that were not significant?
Bonferroni’s correction procedure assumes that the alpha level should be adjusted downward to consider chance capitalization. All the results in Table 4 were not significant even with Bonferroni’s correction because the new cut-off for 15 comparisons is 0.003 and not 0.05 as being usual.

The term “almost perfect” at the start of the discussion section is subjective, and probably should be replaced with “no statistical difference”.
We disagree with the reviewer. Since that we used the definition of Landis and Koch, which can be seen in the Reference section to compare both kinds of questionnaires, we decided to apply the strength of agreement definition in the reference previously mentioned.

The term “a friendly interface” is also subjective; it is not supported by anything in the paper, and is a term that tends to be frowned on by those doing research in human factors. The change has been made. Now, the sentence is: “…the high agreement could be explained by the use of a set of questions with a pre-defined menu of alternatives as a part of the program”.

The description of the authors’ tool as an open source program on pages 9 and 10 is not technically correct. On page 9, the term “Open Source” has been deleted. On page 10, the change has been made. Now, the sentence is: “…the usefulness of using an Open Source program in a PDA to collect data about sexual behavior…”

The statement that the authors’ software “demonstrated a low rate of errors” seems to be subjective and not supported by their data”. The statement has been modified. Now, the sentence is: “Another advantage of the system used is that it demonstrated a low rate of inconsistencies and missing values”, which are consistent with the data.

In Table 1, listing the percentage of females seems to be redundant, since we know the percentage of males. The same for listing the percentage of those who completed high school or higher, and those in the high income category. Also, the definition of high income is missing. The changes have been made. Also, definitions of low educational level and low income were added in the section Methods.

**Minor Essential Revisions:**

*The last full paragraph on page 9 seems to be introductory information that might fit better in the background section.*

This paragraph has been deleted.

*The term “Fox Pro” is actually one word, “FoxPro”*

This change has been made.

**Discretionary revisions:**

*The words “firstly” and “secondly” at the bottom of page 9 might be more correct as “first” and “second”.*

This paragraph has been restructured; some parts have been changed based on the comments of the other reviewer.

*The word “contrinutions” on page 11 should likely be changed to “contributions”.*

The change has been made.

*The term “non responses” on page 3 should likely be hyphenated.*

The change has been made.
In accordance with the corrections and suggestions of reviewers, we look forward to and appreciate your consideration for publication. Please feel free to contact Antonio Bernabe-Ortiz, MD, with any questions, at (511) 9506-0550 or email 03887@upch.edu.pe or abernabeo@gmail.com.

Sincerely,

Antonio Bernabe-Ortiz, MD.  
Research Assistant  
Universidad Peruana Cayetano Heredia.