PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Psychometric properties of the WHO Violence Against Women instrument in a female population-based sample in Sweden
AUTHORS	Nybergh, Lotta; Taft, Charles; Krantz, Gunilla

VERSION 1 - REVIEW

REVIEWER	Dr Holly Johnson Associate Professor Department of Criminology University of Ottawa Ottawa, Canada K1N 6N5 I declare that I have no competing interests.
REVIEW RETURNED	05-Oct-2012

GENERAL COMMENTS	My comments are very minor. Overall, this is a well-written article about an important methodological question in intimate partner violence research: can instruments designed to measure women's experiences of partner violence be applied with equal validity to male populations. The Methods section should specify that it was a postal survey (this is in the abstract only). One suggested reference for review of literature regarding gender differences in IPV on page 4: Ansara & Hindin (2010) Exploring gender differences in the patterns of intimate partner violence in Canada: a latent class approach. J Epidemiol Community Health. 64:849-854. It would be good to know how representative the sample is according to the socio-demographic factors presented in Table 1 according to census or some other data source. Also, the response categories for Educational Level do not appear to be exhaustive, ie., what about other types of training and certification that is not acquired in university? Specify how the WHO guidelines for ethical conduct were addressed on a postal survey, eg, confidentiality, safety for respondents, minimizing and responding to emotional trauma. Additional limitations could be mentioned: (1) the sub-sample of respondents who answered both the VAWI and the NorAQ is relatively small; (2) violence experienced earlier in the lifetime may have been undercounted because of the structure of the questionnaire (this is mentioned on page 16 but merits mention as a limitation); and (3) question wording on the NorAQ may threaten the reliability of estimates of psychological violence, ie., "systematically"

	and "for a longer period" may not consistently have the same meaning for all study participants, unless definitions were provided that are not shown in the article. Or, is it possible this is a reflection of the translation from Swedish to English? On page 19, there is a typo in the 3rd line of physical violence: "trashed" I think should be "thrashed". With respect to the companion article, "Psychometric properties of the WHO Violence Against Women instrument in a female population-based sample in Sweden" I also recommend it for publication. The suggestions above also apply to this article in addition to a few others: Is it a requirement for sample selection that respondents are currently in a relationship or have had a relationship in the past? This is not specified in the Methods section; however, on page 9 it states "the rest of the sample was single, widowed or divorced, but had previously been in a relationship". This should be clarified. Can you explain why the lifetime prevalence of IPV among the sub-sample in Table 5 is so different than the prevalence shown in Table 4? Is it an artefact of sampling? On page 6, 5th line of 2nd paragraph, "higher among men" should be "higher among women".

REVIEWER	Ana Bernarda Ludermir. Associate Professor. Universidade Federal de Pernambuco, Brazil. Conflicts of interest I have no potential conflicts of interest, including specific financial interests and relationships and affiliations relevant to the subject matter or materials discussed in the manuscript.
REVIEW RETURNED	12-Nov-2012

THE STUDY	This paper draws on a survey carried out in Sweden to investigate psychometric properties of the WHO Violence Against Women instrument. The psychometric properties of the VAWI have been documented at least in three papers (Garcia-Moreno et al., 2006; Ellsberg et al., 2008; Schraiber et al., 2010), both in high (Japan) and lower-middle-income countries (Ethiopia and Brazil, among others). The title implies that the paper's key objective is to investigate the psychometric properties of the VAWI. However, to be used in Sweden, the questionnaire should be translated and independently back-translated and discussed to establish accuracy, cognitive understanding, and cultural acceptability (Garcia-Moreno et al., 2006). Furthermore the instrument under test (VAWI) may be compared to other considered as the gold standard and sensitivity, specificity, positive predictive value, negative predictive value and the kappa calculated. The NorAQ is not commonly used and was developed for investigations and comparisons in the Nordic countries of various forms of violence, including different perpetrators, as well as abuse in the health care system, not specifically for intimate partner violence. Ellsberg M, Jansen HAFM, Heise, L, Watts, CH, Garcia-Moreno C.

	Intimate partner violence and women's physical and mental health in the WHO multi-country study on women's health and domestic violence: an observational study. Lancet 37, pp. 1165-1172, 2008.
RESULTS & CONCLUSIONS	There was considerable disagreement between the prevalence found by the two instruments. However, these were based in small figures. For example, only 2 women reported psychological violence in the NorAQ (Table 5). Why the number of women presented in Table 2 is 534 and not 573? The authors should present a table with the association between intimate partner violence and self-rated health and having witnessed parental (or equivalent) physical violence. Regarding interpretation, authors should mention the selection bias related to the fact that data collection was based on a mailed self-report questionnaire. Finally, as they assumed “the sample size used in this comparison prohibits any strong conclusions”. A larger study and with better response rates would be necessary in order to verify the research question.

REVIEWER	Dr Ligia Kiss lecturer London School of Hygiene and Tropical Medicine UK No conflict of interests.
REVIEW RETURNED	29-Nov-2012

REPORTING & ETHICS	The ethical and methodological implications of relying on self-completed questionnaires versus face-to-face interviews are not clear and should be further explored by the authors. The WHO multi-country research and related ethical and safety recommendations focus exclusively in face-to-face interviews. Issues related to confidentiality and safety of respondents may pose some extra challenges in a postal survey that should be addressed by the research team. For example, because of increased risk of retaliatory violence by a partner in case of participation in the survey, it is possible that women experiencing IPV may be more reluctant to answer the questionnaire or disclose violence if not assisted by a trained interviewer. Moreover, the authors should details how the research was presented to interviewees and referrals to sources of support were conducted.
GENERAL COMMENTS	This paper assesses the validity of the WHO violence against women instrument in a population sample in Sweden. This instrument has widespread use in the field and, to date, there has been limited research on its psychometric properties. This makes this a timely and much needed study. However, the study has important methodological limitations as discussed below. The main limitation of this study was the low response rate (62%). The drop-out rate was higher among young, unmarried, lower income and foreign women. It is unclear how this limitation biased the sample and affected the results, especially considering that some of these factors have been previously identified as predictors of intimate partner violence (age, marital status, income). Conclusions from the results of the comparative analysis between the WHO instrument and the NorAQ were limited because of the small sample size for this analysis (n=77). A second data collection

	was conducted for this comparative component, including the recruitment of 20% of the original sample. However, the response rates were again low which, added to the low response rates in the first round, is very likely to have resulted in important selection bias. It would be helpful to include a table comparing socio demographic indicators and factors associated to intimate partner violence (IPV) between these two samples. Additionally, confidence intervals should be reported in table 5. In table 2, page 11, items in the instrument should be described instead of just indicated by numbers. In page 14, lines 26-39, the authors state that 'physical and sexual violence are more likely to occur in conjunction' and 'psychological violence may occur in isolation'. It is not clear in the paper what the basis for these conclusions was. It would be helpful if the authors presented statistical evidence from the study to support these conclusions. Despite limitations, this article introduces an important and timely discussion in the field. I recommend publication after revision.

VERSION 1 – AUTHOR RESPONSE

Reviewer 1:

Some of our responses will overlap those we gave to your comments on the companion article.

The Methods section should specify that it was a postal survey (this is in the abstract only).

Authors' response: Thank you for pointing this out. The following has been added to the methods section, under "Procedure, study population and response rate": "Data was collected by means of a postal survey between January and March 2009."

One suggested reference for review of literature regarding gender differences in IPV on page 4: Ansara & Hindin (2010) Exploring gender differences in the patterns of intimate partner violence in Canada: a latent class approach. J Epidemiol Community Health. 64:849-854.

Authors' response: This comment was probably aimed at the companion paper, where gender differences in IPV are discussed in the introduction. The current paper's introduction focuses on the Violence Against Women Instrument (VAWI) instead.

It would be good to know how representative the sample is according to the socio-demographic factors presented in Table 1 according to census or some other data source. Also, the response categories for Educational Level do not appear to be exhaustive, ie., what about other types of training and certification that is not acquired in university?

Authors' response: Due to financial and time constraints we are unfortunately not able to conduct an exact comparison between the socio-demographic factors in Table 1 and the Swedish population of women in 2009. However, we were able to retrieve information and compare our final sample with the Swedish population of women in 2009 with respect to 1) the five age groups and 2) civil status married/unmarried/divorced or widowed). This comparison showed that those of younger age (18-29 years) are underrepresented in our final sample. Our sample was representative for civil status. We also compared 3) country of birth (Sweden/outside Sweden), 4) income and 5) educational level, although this comparison was somewhat hampered by differing age groups and categorizations

between data available from the national population register maintained by Statistics Sweden and the data used in our study. The results showed that those born outside Sweden, those with a high school degree and those with a high annual income (320 000 Swedish kronor or above) were underrepresented in our final sample. Data for and analysis of this comparison may be requested from the corresponding author, and we have included a Table of this comparison for the referee. Finally, we assume that the referee is referring to vocational training and perhaps associate degrees. Vocational training programs are incorporated into Swedish high schools and curriculums are similar to college prep curriculums in that vocational students meet eligibility requirements for admission to university. In other words, there are no clear differences between the two and we therefore included both in the “high school” category. Also, there is no counterpart to community colleges in Sweden. Hence, we feel that the educational categories used in our study are meaningful within a Swedish context.

Specify how the WHO guidelines for ethical conduct were addressed on a postal survey, eg, confidentiality, safety for respondents, minimizing and responding to emotional trauma.

Authors’ response: The following has been added under “ethical considerations”: “The Regional Ethics Review Board located in Gothenburg gave approval for this study (Dnr: 527-08) and the WHO ethical and safety recommendations for research on domestic violence against women as applicable to a postal survey were followed.(29) For example, a letter was sent to prospective respondents in advance to inform them about the upcoming survey; this provided them with the opportunity to decline the survey before receiving it. Also, although the sampling frame was based on registered individuals, only one survey per household was sent for ethical and safety reasons. Additionally, full anonymity and confidentiality were guaranteed and contact information to a general practitioner (third author on this study), a psychologist and a contact person at Statistics Sweden was provided for additional information and/or referral. The survey was entitled “A study on conflicts, relationships and health”. The study description that followed the title stated that the study assesses IPV.”

Additional limitations could be mentioned: (1) the sub-sample of respondents who answered both the VAWI and the NorAQ is relatively small; (2) violence experienced earlier in the lifetime may have been undercounted because of the structure of the questionnaire (this is mentioned on page 16 but merits mention as a limitation); and (3) question wording on the NorAQ may threaten the reliability of estimates of psychological violence, ie., "systematically" and "for a longer period" may not consistently have the same meaning for all study participants, unless definitions were provided that are not shown in the article. Or, is it possible this is a reflection of the translation from Swedish to English?

Authors’ response: We agree with points 1-2 and have added the following to the limitations section: “Furthermore, the earlier-in-life estimates may have been underestimated due to a minor detail on the questionnaire lay-out.” As well as: “Finally, the sub-sample of respondents who answered both the VAWI and the NorAQ is small, which limits our ability to draw conclusions or generalize to the target population.“ Point 3 is interesting; however, NorAQ has been shown to be valid and reliable in previous studies by other authors. As it was not the aim of the current study to investigate the validity of NorAQ, we consider further exploration of this point to be outside the scope of this study.

On page 19, there is a typo in the 3rd line of physical violence: "trashed" I think should be "thrashed".

Authors’ response: Thank you for noticing this; we have corrected the spelling.

Is it a requirement for sample selection that respondents are currently in a relationship or have had a relationship in the past? This is not specified in the Methods section; however, on page 9 it states "the rest of the sample was single, widowed or divorced, but had previously been in a relationship". This should be clarified.

Authors' response: We have now clarified this in the methods section: "A requirement for the sample selection was that the respondent was currently or had previously been in an intimate relationship." However, one woman (and one man in the companion article) responded that she had not been in an intimate relationship, although she later disclosed psychological violence by an intimate partner. It is possible that she misunderstood the screening question regarding partner status, that she thought about another perpetrator than an intimate partner when reporting violence, or that she defined an intimate relationship differently during the different sections of the survey. Since we are unable to know the reasons behind the discrepancy, we decided to include her in the final sample. In our opinion and in line with the WHO ethical and safety recommendations for research on domestic violence against women, it is important not to neglect data on IPV and to take cases seriously when a respondent has trusted us with information of abuse.

Can you explain why the lifetime prevalence of IPV among the sub-sample in Table 5 is so different than the prevalence shown in Table 4? Is it an artefact of sampling?

Authors' response: This is a very relevant question. The difference stems partly from the fact that Table 5 presents life-time prevalence whereas Table 4 presents past-year and earlier-in-life prevalence separately. However, even when life-time prevalence is calculated for the total final sample, there is still a significant difference for psychological IPV (37.5% vs. 17.1%) and physical IPV (18.7% vs. 6.8%), whereas sexual IPV is similar (10.7% vs. 9.3%). We do indeed believe that this is an artefact of sampling. The sample used for Table 5 is small and yields uncertain estimates. As per your earlier comment, we have added a sentence on this in the limitations section (see above).

On page 6, 5th line of 2nd paragraph, "higher among men" should be "higher among women".

Authors' response: Thank you very much, this has been corrected.

Reviewer 2

The psychometric properties of the VAWI have been documented at least in three papers (Garcia-Moreno et al., 2006; Ellsberg et al., 2008; Schraiber et al., 2010), both in high (Japan) and lower-middle-income countries (Ethiopia and Brazil, among others).

Authors' response: The Schraiber et al. study from 2010 has indeed been an important reference for the current manuscript as it is, to the best of our knowledge, the only published paper on psychometric properties assessing aspects of validity of the VAWI. We also refer to the Garcia-Moreno et al paper from 2006 in our introduction, as their paper publishes a combined Cronbach's alpha for all sites included in the WHO Multi-Country study as a measure of internal reliability. However, the paper does not include information on aspects of the instrument's validity. Finally, we were not able to find information on the psychometric properties of the VAWI in the Ellsberg et al article from 2008. The article mentions that since "the instrument has not previously been validated in the study countries and the mean scores varied widely in the sites, in each site, significance was measured by use of negative binomial-regression techniques." (p. 1168). However, we have revised the following sentence in the introduction: "Extensive pre-testing, independent back-translations and piloting of the questionnaire were conducted.(9)"

The title implies that the paper's key objective is to investigate the psychometric properties of the VAWI. However, to be used in Sweden, the questionnaire should be translated and independently back-translated and discussed to establish accuracy, cognitive understanding, and cultural acceptability (Garcia-Moreno et al., 2006).

Authors' response: We have revised the following sentence in the methods section, "assessment instruments": "The VAWI items were translated and adapted to a Swedish context by a senior researcher (third author) with extensive knowledge about intimate partner violence." An independent back-translation was not conducted, although we agree that future surveys in Sweden using the VAWI should also include this step.

Furthermore the instrument under test (VAWI) may be compared to other considered as the gold standard and sensitivity, specificity, positive predictive value, negative predictive value and the kappa calculated. The NorAQ is not commonly used and was developed for investigations and comparisons in the Nordic countries of various forms of violence, including different perpetrators, as well as abuse in the health care system, not specifically for intimate partner violence.

Ellsberg M, Jansen HAFM, Heise, L, Watts, CH, Garcia-Moreno C. Intimate partner violence and women's physical and mental health in the WHO multi-country study on women's health and domestic violence: an observational study. Lancet 37, pp. 1165-1172, 2008.

Authors' response: Thank you for this comment; we agree that sensitivity, specificity, and positive and negative predictive values are good examples of psychometric aspects to be considered by future studies. On the one hand, we could have used NorAQ for this purpose. However, the two questionnaires are indeed not identical and differing results were therefore expected from start. Hence we did not include NorAQ as a golden standard per se to judge which questionnaire would be better suited for assessing partner violence. Instead, we wanted to see if the prevalence rates obtained from the two instruments could be compared in a meaningful way. NorAQ was chosen since it is the only available, validated instrument assessing violence among both a female and a male (see companion article) population-based sample in a Swedish context. Only the NorAQ violence items applicable to an intimate partnership were included in our study. Moreover, there is evidence for its good validity and reliability. Finally, as we were inspired by the Schraiber et al study for cross-cultural comparisons of the VAWI's psychometric properties, the selected analyses are similar to those explored by Schraiber et al (such as Cronbachs alpha and exploratory factor analysis). However, this question also raises a very important and challenging point about the difficulties in defining a gold standard within IPV research, since there is no objective diagnostic test of IPV in the same way as for diabetes, for example. Instead, different answers and prevalence rates will be obtained depending on the questions asked, the framing of the survey and so on – what, then, should be considered the "true" gold standard? These difficulties also illuminate the need for standardized instruments for comparisons of IPV exposure between and within countries.

There was considerable disagreement between the prevalence found by the two instruments. However, these were based in small figures. For example, only 2 women reported psychological violence in the NorAQ (Table 5).

Authors' response: We agree with this point: the sample used for Table 5 is small and yields uncertain estimates. In line with similar comments from the other reviewers as well, we have added a sentence on this in the limitations section: "Finally, the sub-sample of respondents who answered both the VAWI and the NorAQ is small, which limits our ability to draw conclusions or generalize to the target population." This is also why our conclusion builds on the other analyses included in the study and omits this comparison. Moreover, only the difference for psychological IPV was statistically significant, which we further investigated on an item-based level: "only the difference for psychological IPV was statistically significant (17.1% vs. 2.6%; $p<0.05$). This difference owed principally to the VAWI items "Insulted me in a way that made me feel bad about myself" (16.9%), for which NorAQ has no corresponding item, and "Belittled and humiliated me in front of other people" (6.5%). Prevalence rates for the two other items on this scale were similar to corresponding items in the NorAQ (see

Appendix 1)."

Why the number of women presented in Table 2 is 534 and not 573?

Authors' response: Since the principal components analysis uses list-wise deletion for handling missing data (as indicated in the footnote of Table 2), the analysis results in a smaller amount of respondents. Hence, the PCA used for the current study is based on those 573 respondents with no missing values, which amounts to 534 women.

The authors should present a table with the association between intimate partner violence and self-rated health and having witnessed parental (or equivalent) physical violence.

Authors' response: Thank you for this comment; we agree that it is desirable to give a point estimate for the known groups analysis. Since we have already reached the allowed maximum number of tables for a research article in BMJOpen, we have given the results in the text instead, and added the following: "Specifically, a significantly larger proportion of respondents who reported exposure to violence also reported worse health (Chi-Square (1, N=573) = 26.1, $p<0.05$) and having witnessed parental physical violence (Chi-Square (1, N=573) = 11.5, $p<0.05$) than did those not reporting exposure."

Regarding interpretation, authors should mention the selection bias related to the fact that data collection was based on a mailed self-report questionnaire.

Authors' response: We have added the following under "methodological considerations": "The VAWI was designed for and is primarily used in face-to-face interviews(5), whereas the current study administered the VAWI via a postal survey. The implications of different modes of data collection are difficult to assess due to multiple influencing factors, including the method of initial contact with the respondents, visual versus oral presentation of response choices, method of sampling as well as differing cultural and social contexts.(33) Previous studies have found disclosure of sensitive topics to be higher in self-administered modes compared to face-to-face interviews,(33) also when assessing IPV(34). However, there is a scarcity of experimental or randomized study designs comparing different modes of data collection.(33)

Nonetheless, the main known limitation of postal surveys is lowered response rates.(34) The current study included two reminders in an effort to minimize drop-out rates. Non-responders were over-represented by…"

The added references are:
33. Bowling A. Mode of questionnaire administration can have serious effects on data quality. Journal of public health. 2005;27(3):281-91.
34. Walby S. Improving the statistics on violence against women. Statistical Journal of the United Nations Economic Commission for Europe. 2005;22(3):193-216.

Additionally, the selection bias has been addressed in terms of both the external and internal the drop-out analyses.

Finally, as they assumed "the sample size used in this comparison prohibits any strong conclusions". A larger study and with better response rates would be necessary in order to verify the research question.

Authors' response: We agree that a larger sample is needed for comparisons between the VAWI and NorAQ, and have added a phrase on this in the limitation section (please see above).

Reviewer 3

No conflict of interests.

The ethical and methodological implications of relying on self-completed questionnaires versus face-to-face interviews are not clear and should be further explored by the authors. The WHO multi-country research and related ethical and safety recommendations focus exclusively in face-to-face interviews. Issues related to confidentiality and safety of respondents may pose some extra challenges in a postal survey that should be addressed by the research team. For example, because of increased risk of retaliatory violence by a partner in case of participation in the survey, it is possible that women experiencing IPV may be more reluctant to answer the questionnaire or disclose violence if not assisted by a trained interviewer. Moreover, the authors should details how the research was presented to interviewees and referrals to sources of support were conducted.

Authors' response: Thank you for this interesting reflection. This comment is much in line with a similar comment from reviewer 1 and the following has been added under "ethical considerations": "For example, a letter was sent to prospective respondents in advance to inform them about the upcoming survey; this provided them with the opportunity to decline the survey before receiving it. Also, although the sampling frame was based on registered individuals, only one survey per household was sent for ethical and safety reasons. Additionally, full anonymity and confidentiality were guaranteed and contact information to a general practitioner (third author on this study), a psychologist and a contact person at Statistics Sweden was provided for additional information and/or referral. The survey was entitled "A study on conflicts, relationships and health". The study description that followed the title stated that the study assesses IPV."

Moreover, we have added the following to "methodological considerations": "The VAWI was designed for and is primarily used in face-to-face interviews(5), whereas the current study administered the VAWI via a postal survey. The implications of different modes of data collection are difficult to assess due to multiple influencing factors, including the method of initial contact with the respondents, visual versus oral presentation of response choices, method of sampling as well as differing cultural and social contexts.(33) Studies have found disclosure of sensitive topics to be higher in self-administered modes compared to face-to-face interviews,(33) also when assessing IPV(34). However, there is a scarcity of experimental or randomized study designs comparing different modes of data collection.(33)

Nonetheless, the main known limitation of postal surveys is lowered response rates.(34) The current study included two reminders in an effort to minimize drop-out rates. Non-responders were over-represented by…"

The new references included were:
33. Bowling A. Mode of questionnaire administration can have serious effects on data quality. Journal of public health. 2005;27(3):281-91.
34. Walby S. Improving the statistics on violence against women. Statistical Journal of the United Nations Economic Commission for Europe. 2005;22(3):193-216.

Finally, it would be both resource and time consuming to conduct face-to-face interviews for a population-based sample in Sweden. If interviews were to be conducted over the telephone instead, it would be difficult to know if the perpetrator would be close by when the call is made. A postal survey, on the other hand, may be opened at any time, or thrown away. Furthermore, as most people would

be working during the day-time they would not answer their phones or doors, and in the evening they are occupied with household and other choirs and may, in a Swedish context, be irritated if their private space is invaded during this time. Again, a postal survey is less invasive from this perspective and may be answered at leisure.

This paper assesses the validity of the WHO violence against women instrument in a population sample in Sweden. This instrument has widespread use in the field and, to date, there has been limited research on its psychometric properties. This makes this a timely and much needed study. However, the study has important methodological limitations as discussed below.

The main limitation of this study was the low response rate (62%). The drop-out rate was higher among young, unmarried, lower income and foreign women. It is unclear how this limitation biased the sample and affected the results, especially considering that some of these factors have been previously identified as predictors of intimate partner violence (age, marital status, income).

Authors' response: We agree with this point and declining response rates is unfortunately a growing concern for survey-based studies in Sweden and the Nordic countries in general. While the drop-out rates in the current study may be considered high, they are, however, in line with other survey-based studies in the Nordic countries today. For example, the response rate for a population-based study assessing IPV by postal surveys in Finland in 2005 and to which we compare our prevalence rates in the manuscript, was 62% (N=4464) compared to 62% in the current study. The authors conducted a similar study in 1997 when the response rate was 70% (N=4955), which points at the declining trend. We also agree with the point that IPV has probably been under-reported due to the socio-demographic characteristics of the respondents. Hence the manuscript includes the following under "methodological considerations": "Moreover, non-responders were over-represented by young and unmarried women, women with lower income and by those born outside of Sweden. Exposure rates to IPV have been found to be especially high in these groups,(21, 25) which may further contribute to under-estimated prevalence rates and less robust component solutions in our study." It could be added that the strength of the associations found in the known-groups analysis are therefore also probably under- rather then overestimated, which would further strengthen the evidence found from this analysis.

Conclusions from the results of the comparative analysis between the WHO instrument and the NorAQ were limited because of the small sample size for this analysis (n=77). A second data collection was conducted for this comparative component, including the recruitment of 20% of the original sample. However, the response rates were again low which, added to the low response rates in the first round, is very likely to have resulted in important selection bias. It would be helpful to include a table comparing socio demographic indicators and factors associated to intimate partner violence (IPV) between these two samples.

Authors' response: The random sample taken for the second data collection was drawn from the final sample of the first data collection. In this regard the idea was that they would have similar socio-demographic characteristics. However, since the final sample of the second data collection was small as the referee correctly points out, we do not claim that the socio-demographic characteristics are generalizable to a population-based level and are cautious in drawing conclusions from this comparison. The following has been added to the limitations section: "Furthermore, the sub-sample of respondents who answered both the VAWI and the NorAQ is small, which limits our ability to draw conclusions or generalize to the target population." Finally, we did conduct this comparison for the referee by use of a two-proportion z-test with Bonferroni adjustment. No statistically significant differences were found (Table has been included for the referee).

Additionally, confidence intervals should be reported in table 5.

Authors' response: Confidence intervals are not reported for table 5 since the estimates are unstable due to the low number of respondents. We do agree with all reviewers on the uncertainty of estimates for this analysis and have therefore added the following to the limitations section: "Finally, the sub-sample of respondents who answered both the VAWI and the NorAQ is small, which limits our ability to draw conclusions or generalize to the target population." This is also why our conclusion builds on the other analyses included in the study and omits this comparison.

In table 2, page 11, items in the instrument should be described instead of just indicated by numbers.

Authors' response: We have revised this point, and the items have now been described in Table 2.

In page 14, lines 26-39, the authors state that 'physical and sexual violence are more likely to occur in conjunction' and 'psychological violence may occur in isolation'. It is not clear in the paper what the basis for these conclusions was. It would be helpful if the authors presented statistical evidence from the study to support these conclusions.

Authors' response: Thank you for this pertinent comment. The following has been revised under the discussion section related to internal validity: "This solution is understandable in that physical and sexual violence occur to a lesser extent compared to psychological violence, which generally is the most prevalent form of IPV.(22, 31)" Finally, after revisiting the Romans et al reference we noticed that the authors of that study combined emotional and financial violence into one category, whereas we assess psychological violence alone. We therefore included an additional reference which also uses the VAWI and presents Venn Diagrams. In this article (see below) and in similar lines with the Romans et al study, psychological violence is the most frequently occurring form of violence (25.4%), compared to both physical (0.6%) and sexual (0.1%) violence as well as with the overlapping forms of IPV (ranging from 6.5% to 0.1%; past-year prevalence).

The added reference is:
Vung N, Ostergren P, Krantz G. Intimate partner violence against women in rural Vietnam- different socio-demographic factors are associated with different forms of violence: Need for new intervention guidelines? BMC Public Health. 2008;8(1):55.

Despite limitations, this article introduces an important and timely discussion in the field. I recommend publication after revision.

Additionally, a mistake was corrected in table 1: the correct percentage of women in heterosexual relationships is 83.2 (N=477) and not 98.8 (N=566) as previously indicated. We do not know whether the women who are single, widowed or divorced in our sample have previously been in same-sex or heterosexual relationships. Finally, a clearer reference to the companion paper was made on page seven by including the name of the title.