Author's response to reviews

Title: The Buffering Effect of Relationship Satisfaction on Emotional Distress in Couples

Authors:

Gun-Mette B. Røsand (gbro@fhi.no)
Kari Slinning (kari.slinning@r-bup.no)
Malin Eberhard-Gran (malin.eberhard-gran@fhi.no)
Espen Røysamb (espen.roysamb@psykologi.uio.no)
Kristian Tambs (kristian.tambs@fhi.no)

Version: 2 Date: 23 September 2011

Author's response to reviews: see over
Dear Dr. Jimmar Dizon

Please find enclosed the revised manuscript:

The Buffering Effect of Relationship Satisfaction on Emotional Distress in Couples
Gun-Mette B. Røsand, Kari Slinning, Malin Eberhard-Gran, Espen Røysamb and
Kristian Tambs

Thank you for the reviews and for your comments on our manuscript. We are pleased
that you and the reviewers found the paper valuable, and that you are willing to
consider a revised draft.

We hope this revised manuscript addresses your concerns. We found the comments
very constructive and useful and believe that their impact have improved the paper
substantially. A detailed description of how we have addressed the issues raised by
you and the reviewers follows below.

Editorial comments:
We have addressed the comments from the reviewers in a revised manuscript and
given a point-by-point response to the concerns.

All changes made when revising the manuscript have been highlighted with ‘tracked
changes’.

Our revised manuscript conforms to the journal style, and the files are formatted in
accordance with BMC Public Health’s Instructions for authors.

On behalf of the authors, sincerely yours

Gun-Mette B. Røsand, PhD
Comments from reviewer 1, Steven Beach

Minor Essential Revisions

1. The description of the sample size and how it ultimately results in the effective sample used in the analyses is somewhat unclear despite substantial attention to describing missing data and the number of data points available. I believe this could be presented in a straightforward way by recapping all steps in the process of creating the final sample on page 14. That is, the authors might say that “90,190 women and 71,648 men were initially recruited to participate in the study, representing XXXXX complete couples. However, for 6542 women and 6586 men it was not possible to impute one or more key study variables resulting in a final, usable sample of 83,648 women and 65,062 men who comprised 62,956 usable couples”.

As a side note, it appears that the percentages given at the top page 14 are incorrect (reversed for men and women). The percentage missing should be larger for men than for women because the denominator is smaller and numerators are similar.

Our response:
We are grateful for this comment. We have now included the suggested sentence in the text (p.14), and deleted the original sentence. We have also elaborated the description of the sample in the “Participants” paragraph (p. 9) trying to make it even clearer.

The reviewer is correct regarding the percentages given at page 14. The percentage missing is 7.3% for women, and 9.2% for men. We have included the percentages at page 14. We thank the reviewer for making us aware of this error.

2. The analyses reported in tables 1 and 2 would probably be more informative if (the) one of the columns reported a standardized Beta weight for each variable, allowing direct comparisons across predictors. Currently the authors do not clearly define what they mean by “crude b” and “adjusted b” and this makes the table less clear than it could be. Prior to the current table 1 and table 2 it would be appropriate to provide a correlation table for all major variables in the analyses as well as the means and standard deviations for non-dichotomous variables.

Our response:
SCL-5 was standardized; therefore the non-standardized regression coefficient (b) shows the expected difference in standard deviations in SCL-5 per predictor scale unit, which (for dichotomous variables) is the same as Cohens d. In most cases, the independent variables in the analyses were also standardized, implying that the b estimates are identical to the β estimates. For dichotomous variables (not standardized) the β estimates are computed under the table, to make the effect sizes comparable. We have added a sentence in the text on p.16 to make this clear to the reader. We have also added one sentence under Table 2 and Table 3 to clarify this issue.
**Crude b and adjusted b:**
Crude b = the association between the dependent variable and each of the risk factors, without controlling for the other variables.
Adjusted b = the associations between emotional distress and each of the risk factors after mutually controlling for all variables in the regression analyses. That is, the unique effects of each of the independent variables.
We have added one sentence under Table 2 and Table 3 to clarify that all adjusted effects are adjusted for all other predictors.

**Correlation table prior to the current tables:**
We have now included a table showing Pearson correlations between all the predictors in the analyses. The means and standard deviations for the independent variables (unstandardized) are also presented in the table (Table 1).

3. I would recommend against calling out the tables 3-6 all at once on page 16. The simplest solution may be to simply drop the section with the sub-heading “Interaction effects” and move on to the next section.

**Our response:**
We agree with the reviewer. We have deleted the section with the sub-heading “Interaction effects” as suggested (p. 17).

4. On page 18 I would recommend against including RS as one of the risk factors included in the calculation of overall impact of partner relationship satisfaction. Including own RS makes the results presented in the two figures non-comparable.

**Our response:**
The two figures are not directly comparable, and that is not the intention. The intention is to show the buffering effect of own RS and partner’s RS respectively, against the total effect of all the other risk factors in the study. Therefore, we prefer to retain the figures as they are.

We have deleted the last sentence in the Results section to avoid possible confusion.

**Discretionary revisions:**
On page 7 – use “buffering hypothesis” and “buffering effect” rather than “buffer hypothesis” and “buffer effect”. Likewise on page 19, use “buffering effect” rather than “buffer effect”.

**Our response: This is corrected on p. 7 and p.20, likewise on p.23).**

On page 13, the description of the use of mean substitution is either not adequately described or is not persuasive. It is not customary to use mean substitution to replace missing values because of its effect on reducing standard deviations around the
resulting grand mean. The authors should explain why this was not a problem in this case.

Our response:
We did not use regular mean substitution to impute values for missing scores on female and male income and education. See the text p.14: “Here, neither EM imputation nor regular mean substitution appeared to be suitable. Instead mean values on valid demographic data for non-responders compared to responders for a particular variable were used to choose a suitable constant for replacing missing values for female and male income and education”.

However, the reviewer’s comment may reflect that the description of the procedure used to replace missing values on these variables was unclear. We have therefore partly re-written the last part of the “Treatment of missing values” paragraph. We hope this makes the description clearer.

We acknowledge that these replacements of missing data change the variation for the education and income variables a little. As explained in the revised manuscript, we used two different constants for replacement of missing education data. There was one category for “other education” that was recoded 3.5 for women and 4 for men. Blanks were recoded to the lowest category. This recoding increased the standard deviations, from 1.44 to 1.46 for men, and from 1.28 to 1.36 for women. Standard deviation for family income changed from 2.20 before missing replacement to 2.18 after missing replacement.

Because the differences in standard deviations before and after imputation were small or moderate, we don’t think these replacements have caused an important bias.

On page 17 the sentence “All, except from one, significant effects....” should be “All except one of the significant effects....”

Our response: This is corrected (p. 18).

Comments from reviewer 2, Joseph M Trombello

Compulsory:

1: I’d like to see a much stronger case made for why the authors have chosen to focus on first-time motherhood. This transition in the first paragraph of “other risk factors associated with symptoms of depression” seems a bit jarring to what has come before. Furthermore, their comment in the next paragraph that “the risk factors for depression in pregnant women seem to be very much the same as for women in general” seems to argue against the need to focus on pregnant women as a special group. Why have the authors chosen to focus on this sub-sample? How has this been an understudied area? Why should the reader care? How would focusing on
this sample of women help to advance the field or theories about the relationship between marital satisfaction and distress/depression?

This concern is echoed in their concluding paragraph that says “based on earlier research on the association between relationship satisfaction and emotional distress, we suppose that these findings are valid for most couples, regardless of pregnancy”. Why, then, have they chosen to focus on pregnant women?

Our response:
First time motherhood was one of the 12 risk factors and was chosen because we hypothesized this could be a risk factor for emotional distress in both men and women in our sample. This assumption is also supported by results from previous research, showing that first-time motherhood is associated with depression during pregnancy. We have added another reference here (p. 5). We have also re-phrased the transition to introduce first-time motherhood as one possible risk factor for emotional distress (p. 5-6).

Regarding why we have chosen to focus on pregnant women:
This study is based on the Norwegian Mother and Child Cohort Study (MoBa). This is a large population based longitudinal study including couples. MoBa provides opportunities for studying factors specific for pregnancy and childbirth, e.g. first-time motherhood, and also for studying general processes and effects. Mothers undergoing their first routine ultrasound examination, performed at gestation week 17-18 were invited to participate together with their male partners. The women, but not the men, were followed up at later times, also after pregnancy. Our aim was to study couples, and because we only had data from the male partners at gestational week 17-18 in pregnancy, we had to use couples who expected children.

Nevertheless, we think there are reasons to believe that the results may be generalized to couples in other phases of life. The fact that the risk factors for depression in pregnant women seem to be very much the same as for women and men in general supports this suggestion.

We have added some text in the “Strenghts and Limitations” section about the extent to which we may generalize from pregnant couples to couples in general.

The association between relationship dissatisfaction and depressive symptoms in couples is not an understudied area, and that is not our assertion either. However, to date, most research on relationship dissatisfaction and psychological distress has been based on small samples. We believe that the MoBa sample is unique and provides the opportunities to obtain precise estimates and to investigate interaction effects that are difficult to establish in smaller samples.

2: I would also like to see a stronger case for why the authors selected their 12 risk factors? It feels more like that they were chosen because they are in the available dataset than for any substantive or theoretical reason.
Our response:
The risk factors were chosen because we hypothesized they would be of importance for emotional distress in men and women in our sample. Our assumptions were based on results from previous studies. The risk factors were also chosen simply because they were among the available variables in the data-set.

The main aim was however not to investigate main effects of the risk factors, but to study buffering effects of own and partner’s relationship satisfaction (c.f. the title of the manuscript). We considered the risk factors to be the most appropriate for the main purpose of the study. The risk factors chosen were a possible range of variables usable for testing buffering effects of relationship satisfaction on emotional distress. According to our hypotheses/assumptions, these risk factors were representing certain strains that could be moderated by relationship satisfaction. We have added some text to underline that investigating buffering effects was the most important aim of the study (p. 2, 8, 20).

Even if we regarded the buffering effects (Table 4-7) as the most important results, we think it is reasonable to present the main effects first.

3: Are there differences between fathers who completed the full 5 relationship items and the ones who completed less than this number? I’m just worried about data imputation a bit given that fathers could have completed a variety of the already very limited number of items in this scale. On the other hand, the alpha values and correlation numbers are helpful.

Our response:
We did not generate imputed values for men with less than five valid items, regardless of whether they completed the 10- or 5-item version. In other words, we did not impute the data from the 5-item version (see p.13).

We have tested for interaction between RS score and a variable indicating if the men had completed a 10-item or a 5-item RS version. The results from this study are the same regardless of whether the men had the RS5 or the RS10. This information is also included in the Results section (p.19).

4: This cannot be changed, but the single item social support measure is less than desirable.

Our response: We agree with the reviewer. This is now mentioned under limitations (p.24).

5: Why were women given a checklist of 53 somatic disease items while men only had 19 such items? Could this have affected their results?

Our response:
The women in MoBa are followed up at many time points (eight so far), while the men have only received one questionnaire. The competition for space in the male questionnaire was therefore hard, and only the illnesses judged to be the most important ones were included.

It was not possible to generate exactly the same disease-variable for men and women because some of the items in the male questionnaire were broader and/or different from those in the female questionnaire. However, we based the final indicators used in the analyses on approximately the same groups of diseases for both sexes; see the text (p. 11 and 12).

Could this have affected the results?

Our response:
This is a good question. There is a danger that the effects would be more washed out for men because of less detailed descriptions of somatic disease. However, the results show the opposite: the main effect of somatic disease on emotional distress is higher for men than women (Crude b: 0.20 for men and 0.17 for women), which doesn’t suggest such a bias.

6. Similar to number 4, this cannot be changed, but the fact that only a very small minority of participants were classified as low relationship quality is a problem for external validity and for clinical significance.

Our response:
We think there is reason to believe that this skewed distribution is not primarily a result of sample selection, but rather reflects a general trend for people to choose the most “favourable” and social acceptable response categories. Previous studies typically find that various well-being and satisfaction measures are positively skewed.

However, there is also reason to believe that most of the men and women in our population are in fact quite satisfied with their relationship. Most couples probably choose to have children when their relationship feels satisfying and stable. Finally the response rate is low, and there may well be some selection towards high relationship satisfaction as well.

We think it is generally accepted that sample selection is less harmful to the estimation of associations than to estimation of prevalences. We have added two sentences about the possible effect of recruitment bias in the Strength and Limitation section.

7: I’d like to see effect sizes computed for their results, especially in tables 1 and 2. Given the huge number of participants, I’m not sure if significant results are due to high power or due to a strong effect. The authors helpfully note this limitation, though, in their opening paragraph of the discussion.
Our response:
The effect sizes are computed as unstandardised regression coefficients (b), with the dependent variable standardized. For predictors which are z-scaled b and β are identical, since the dependent variable is also z-scaled. Beta-coefficients do not depend on “the scaling” of the variables, which makes the various estimates more easy to compare. For the dichotomous variables the b gives (adjusted) difference between the two groups, which is the same as Cohen’s d. For these variables the beta values are reported in the table subtext (Tables 2 and 3). We have also added a sentence in the text (p.16) to make this clear to the reader.
In Tables 2 and 3 the (original) range of the independent variables are given regardless of whether these variables were z-scaled or not. To avoid confusion the title of the column showing the ranges has been changed to “Range (before z-transformation)”.

8: Similarly, given the large number of analyses and comparisons run, I’d like to see some control or correction for multiple comparisons or capitalization on chance. Bonferroni correction may be one such (admittedly conservative) approach.

Our response:
We have chosen a significance level, p < 0.001, which is even lower than the Bonferroni correction with these number of tests. This is true for all the results/ tables, even those in table 4, which is the table with the highest number of analyses.

9: In “the effect of relationship satisfaction on emotional distress” portion of the discussion, the authors note that “this finding has already been demonstrated for women in the present sample and is consistent with previous research on couples”. If so, what is novel or exciting about their finding? Is this merely a replication?

Our response:
The most important aim of this study was to assess whether high relationship satisfaction protects against severe effects of certain stressful events, for both men and women. Although we found it appropriate to introduce the main effects first, we consider the buffering effects as the most important findings in this manuscript. The large sample size is well suited for testing and precise estimation of interaction effects. Among other interaction effects we were interested in interaction effects between the spouses. Our previous results only included data from women.

Essential:

1: The transition to “buffer effects of relationship factors” needs improving.

Our response:
We have added a sentence in the beginning of the paragraph (p.7).

2: In the “level of relationship satisfaction in couples” section in the discussion, I’d like to see a stronger connection made to the decline in marital satisfaction after pregnancy literature.
Our response: Thank you for this comment. We have added references to studies showing decline in marital satisfaction after pregnancy in the text (p. 20).

Discretionary:

1: The final two sentences of the results are a bit confusing as written.

Our response:
Frankly, we don’t see what was unclear about these sentences. We are probably blind to our lingual shortcomings. We have tried to rewrite the second last sentence in the Results section to make it clearer. We have deleted the last sentence in the section to avoid possible confusion.