Author's response to reviews

Title: Assessment of social psychological determinants of satisfaction with childbirth in a cross-national perspective

Authors:

Wendy W Christiaens (Wendy.Christiaens@ugent.be)
Piet P Bracke (Piet.Bracke@ugent.be)

Version: 2 Date: 14 June 2007

Author's response to reviews: see over
Dear editor and reviewers,

Please find included the point-by-point description of the changes made.

With best wishes,

Wendy Christiaens

Satisfaction with childbirth in Belgium and the Netherlands

New title: Assessment of social psychological determinants of satisfaction with childbirth in a cross-national perspective

<table>
<thead>
<tr>
<th>Reviewer's report:</th>
</tr>
</thead>
<tbody>
<tr>
<td>General</td>
</tr>
<tr>
<td>This is an interesting study, but the report would benefit from substantive editing in the Background, Literature Review, and Results.</td>
</tr>
<tr>
<td>I enjoyed reading the discussion and conclusions sections, which are derived appropriately from the results, and succinctly written.</td>
</tr>
</tbody>
</table>

Major Compulsory Revisions (that the author must respond to before a decision on publication can be reached)

Background
1. There are sweeping, unsupported generalizations in the Background. Try to reference all statements. For example, I think Belgian health care professionals would be rightfully upset by their characterization as providers that treat pregnant women as passive and lacking in knowledge or authority. I think it would be better to describe the two different approaches, with supporting references.

We did not intend to characterize Belgian health care professionals as treating women as passive and lacking in knowledge or authority. This is only a description of what is meant by the medical model. The medical model is an ideal type: this means it is a theoretical, analytical construct that generalizes the conduct of individual actors. ACTION: We reformulated this part of the text to make a clearer distinction between theoretical models and the practice in Belgium and the Netherlands, and we added some references.

2. There is no description of the analytical typology. Use a better source than personal communication for home birth statistics.

OK. ACTION: We added references.

**Literature Review**

3. Labour pain: I would not characterize the relationship between pain and satisfaction as puzzling. The Hodnett reference (i.e. my review) is to a systematic review of the relationship between pain and satisfaction. The single studies referred to in this section are consistent with the findings of the systematic review.

ACTION: Instead of puzzling we wrote: "Reports about the relationship between the intensity of pain and satisfaction provide mixed results”. Some authors report direct relations between pain intensity and satisfaction with childbirth, and others, indirect relations or no relation at all.

4. “Satisfaction” needs to be defined – Either see the Hodnett systematic review for an overview of the complexity of the construct, the various conceptualizations of it, and the problems in measuring it, or be clear about how you have defined it.

We agree. ACTION: We added a new paragraph at the beginning of the literature review.

5. The literature review would benefit from major reorganization. The last paragraph should be the first paragraph; it provides a useful organizing framework, since it is what drove the decisions about content of the survey.

ACTION: We moved the last paragraph to the end of the background section as an introduction to the literature review.

6. At the end of the literature review, please state the purpose of the study. Remove settings from the literature review (see recommendation #10).

ACTION: We added a small paragraph (at the end of the literature review) describing the aim of the study.

**Methods**

8. In the description of the sample, it would help to know how many women gave birth at each hospital and in each midwifery practice, and the percentage of each who completed the questionnaire. This would be preferable to the estimates provided by the midwives and obstetricians.

ACTION: We reorganised the method section (partly) according to the headings suggested by the second reviewer: selection of method, settings, sample size, recruitment and data collection, ethical consideration, measures and data analysis.

Indeed, this would be preferable, but we do not have access to these figures. We asked for these figures in the hospitals, but they could not provide them.

9. The three items in the Expectancy/Experience questionnaire used to measure control had an unacceptably low Cronbach’s alpha, which is not surprising, since the items only measured one aspect of control. But this introduces serious measurement error. Please address this problem.

A distinction should be made between Cronbach’s alpha in clinical situations and Cronbach’s alpha comparing groups. In clinical situations the minimum is 0.90 and 0.95 is desirable. When comparing, groups alphas are quoted ranging from 0.61 to 0.88 (Band and Altman, 1997). An alpha of 0.67 is more satisfactory because we have only three items. The more items, the more likely a high alpha becomes. Cronbach’s alpha gives an indication of the internal consistency of the scale. This means that when we measure several aspects of control, the alpha would decrease. It is because we measure only one aspect of control that we have a good alpha of 0.67.


10. I am confused about the categorization of policlinic and birth clinics as hospital and home, respectively. I think that restricting the analysis to only those who chose hospital or home (and not some in-between variant) would introduce less threat of bias. Alternatively, provide a much more solid justification for their categorization.

Short stay births account for 106 cases. Omitting them from the analysis reduces our sample considerably. In order to preserve the power of the sample, we ran the analysis with place of birth coded as a dummy, consisting of three categories: home, poli and hospital. This results in only two coefficient changes in comparison with our original analysis. First, place of birth (home, as well as poli as hospital) becomes significant in relation to general satisfaction (Home: B = 0.936, p-value = 0.001, CI = [0.381-1.490]; poli: B = 0.852, p-value = 0.003, CI = [0.286-1.417]; Hospital: B = 0.868, p-value = 0.002, CI = [0.309-1.428]). Second, the effect of self-efficacy for satisfaction with the midwife is no longer significant (B = 0.088, p-value = 0.080, CI = [-0.011-0.188]).

Because these changes are minor, we decided to keep the original analysis.
### Results
11. Means should not be reported without their standard deviations.  
12. I am not an expert on hierarchical linear modelling, but the results seem to be overstated, given that many of the relationships are weak. Statistical significance should not be confused with clinical significance.  

OK. We added standard deviations.  
OK. We made weak relations explicit to the reader in the results and discussion section.

### Minor Essential Revisions (such as missing labels on figures, or the wrong use of a term, which the author can be trusted to correct)

1. I do not understand the last sentence of the Background.

   By fitting the model for two countries characterized by distinct maternity care, we want to assess the generalizability of the determinants. We omitted this sentence. We make this point clear elsewhere in the text.

2. In the Literature Review, section on expectations: what is meant by “Expectations refer to a role system”?

   This is sociological jargon: a role system is the set of roles a person fulfils.

### Discretionary Revisions (which the author can choose to ignore)

Unable to decide on acceptance or rejection until the authors have responded. What next?: to the major compulsory revisions

| Level of interest: An article whose findings are important to those with closely related research interests |
| Quality of written English: Needs some language corrections before being published |
| Statistical review: Yes, but I do not feel adequately qualified to assess the statistics. |
| Declaration of competing interests: |
| I declare I have no competing interests. |

### Version: 1 Date: 13 March 2007
Reviewer: Jennifer Fenwick

**Reviewer's report:**

**General**

For Authors: Wendy & Piet

The article is well written from the point of view of level for this journal and readability. I enjoyed reading it. The question posed by the research is well defined. The literature presented is appropriate, relevant and current. Discussion and conclusions are balanced and adequately supported by the data. Limitations are addressed.
Well done and I hope this feedback will be helpful in strengthening your manuscript.

Major Compulsory Revisions (that the author must respond to before a decision on publication can be reached)

Method
The method is appropriate however the paper would be strengthened by stating and justifying the approach taken (For example … The study modelled the relationships between these constructs and childbirth satisfaction using data collected from a self-reported survey…) and by more clearly detailing how the study was conducted.

ACTION: We added a paragraph "selection of method" at the beginning of the method section.

The following are suggestions on how the manuscript could be strengthened;

Sample size
Page 5 – phrase ‘… since population of pregnancy women unknown’… I’m not sure what you mean here. Surely you must have some idea of the numbers of birthing women in each setting? You can also use the national birthing figures to make some comparisons to your sample and make some judgement as to the representative nature or otherwise.

There needs to be some discussion on the numbers needed to power the study

OK
By “… since population of pregnancy women unknown…” we intended to say that there are no lists of pregnant women from which to make a random selection. Women are known or registered from the moment they give birth, but we wanted to administer a postnatal AND a prenatal questionnaire. Therefore a convenience sample was the only way.

We added this to the text, more specifically, the paragraph about sample size.

The sample size \( n \) needed is calculated as:

\[
n = \frac{z^2 \hat{p} \hat{q} - \Delta^2}{\hat{p} \hat{q}}
\]

\[
n = \frac{(1.96)^2 \times 0.5 \times 0.5}{(0.04)^2} = 600.25
\]
Recruitment and data collection

When did the women receive the postnatal questionnaire? Was this sent out with the antenatal questionnaire? Was there a package? How did women know what to do? For example letter with package?

Information sheet? Was there a self addressed prepaid envelope? What I’m trying to get at is that you need to detail what you did so that someone else could replicate the work – there isn’t sufficient detail at present to know exactly how you went about recruiting the women and collecting the data.

Antenatal questionnaire – what did this consist of? Was this another study? Is this when you collected demographic data? What demographic data did you collect? (I know this is stated later but should be clear up front)

Why 30 weeks?

We followed the convention that $\alpha = 0.05$, corresponding with a 0.95 confidence interval, and $|z_\alpha| = 1.96$. If we allow the confidence interval to be 8 percentage points wide, corresponding with a distance $d$ of 4 per cent, the implied sample size $n$ is 600. (The parameter $d$ is the maximum tolerated distance between the proportion $P$ in the population and the proportion $p$ estimated from the sample (Cochran, 1963 p. 71-75).


ACTION: We added this information to the method section and tried to give all required details to replicate the study.

The antenatal questionnaire was part of the same study. A longitudinal design has been applied. Therefore the same concepts were measured before and after birth, hence the antenatal and postnatal questionnaires were similar.

At 30 weeks of pregnancy, plans are being made and expectations are being formed, because the birth is only 10 weeks away. Other studies administered an antenatal questionnaire around the same gestational moment; for example:


More specifically, the W-DEQ was administered at 28-30 weeks gestation by the designer of the scale (Wijma K). We wanted to assess expectations, so this scale was one of the most important scales in our questionnaires. It consists of an antenatal and postnatal version.


The postnatal version of the W-DEQ is administered within two hours after delivery, but the obstetricians and midwives participating in the study judged this undesirable and impossible to get organised. So, the two weeks were chosen for practical reasons. A follow-up some months later would have been great, but there wasn’t sufficient time or budget to realize this more complicated design. Another reason for taking the measurement close to delivery was the memory of pain. Terry and Gijsbers (2000) suggest that women provide 'reconstructed' reports of their labour pain.


---

**Why 2 weeks for postnatal?** – you need to take into consideration the literature that says if you ask women about their childbirth experience in the early postpartum period likely to get a more positive response as ‘glad over’ – and / or ‘grateful to staff’ - women at this stage have not had time to reflect on the experience.

**OK. just read in limitations.**

**State response rates in % as well as n=**

**Page 6 – second paragraph commencing with ‘We are to rely on the goodwill ......’** This whole paragraph is confusing and I’m unsure what you’re trying to say.

**Why was a written consent necessary when you did not collect any identifying or personal details?**

**Measurements**

ACTION: We changed this part of the text to make it clearer. Together with the range between the hospital/midwifery practice with the lowest response rate and the one with the highest, we have given the response rates for Dutch and Belgian hospitals in total.

We want to explain the reason for the difficulties we encountered in calculating the response rates. **ACTION: We rewrote this part of the text.**

The ethical committee asked us to do so.
Pilot on dependent variable (satisfaction with childbirth) work mentioned – great. You might like to make a comment on the internal consistency reliability coefficients (measured using cronbach alpha) and what they mean. For example .. 'pilot testing demonstrated that instrument was valid for the birthing population in both Belgium and the Netherlands …'

Visual Analogue scales (VAS) – a brief comment on the use of these for measuring childbirth labour pain is needed. Who else has used them and how in the childbirth population? In other words you need to demonstrate to reader that this was a reliable and acceptable way to measure pain in the childbearing woman. There is literature that would argue the opposite.

W-DEQ – was developed by who and for what? (I actually know but from what you’ve written your reader will not). This questionnaire ( W-DEQA & B) was designed to measure expectations and experiences – why not use entire scale etc. In other words you need to justify the use of the three personal control items?

We did not use the entire scale because it is developed to measure fear-related emotions related to childbirth. This is too general and shows overlap with our pain measurement. Although the author of the W-DEQ does not mention separate dimensions in the scale, the three control-related items we picked out are grouped together under a separate heading. Assessment of the correlations (Table 3 below) between these items shows consistency.

Self-efficacy measure – Fairly old? example of questions? Results of pilot testing? When translated and tested did results indicate valid and reliable in your birthing population?

We agree about the Pearlin mastery scale being fairly old. However, the scale is widely used within social sciences. It is considered to be a standard scale. Some references from recent studies using the same mastery scale are presented below. Internal consistency is quite good, with a Cronbach’s alpha of 0.79. Construct validity is good since self-efficacy correlates with self-esteem (measured by the Rosenberg self-esteem scale) (r = 0.57) and with the Edinburgh postnatal depression scale (r = 0.4). ACTION: We summed up the items of the scale, and added the Kempen (1992) reference. 

Discussion

Page 14 – 3rd paragraph 3rd last sentence… I’m not sure what you mean by the statement ‘The ambivalent Dutch maternity care…. etc’ and the conclusions you draw. This needs a little more explanation.

This should be clearer in the light of the revised background part. The ambivalence of the Dutch maternity care refers to the two sciences of obstetrics De Vries talks about: one in favour of and one against home births. In fact, we point to the co-existence of the social and medical model in one country. This co-existence creates contradictory expectations, but women cannot fulfil both. In case of a home birth they deviate from the medical model, in case of a hospital birth they deviate from the social model as ideal types.

Any implications for practice?

Prenatal preparation could enhance a mother’s satisfaction with childbirth by providing techniques to maintain control, by the enhancement of self-efficacy and by providing information that establishes realistic expectations about childbirth. We added this to the conclusion.

Minor Essential Revisions (such as missing labels on figures, or the wrong use of a term, which the author can be trusted to correct)

Spacing
Attention to formatting needs addressing
Some track change markings are still apparent in the document

Grammar
Minor editing is needed. When writing you should use ‘past tense’.
Use numerals when 20 and over – below that use words ie., six

References


Reference in 2nd paragraph in Background – ‘personal communication’ needs a name and also should have number assigned.

We changed this to another more solid reference.

**Abstract**

An abstract is needed.

This is caused by a misunderstanding during the upload process. We will include the abstract in the revised manuscript.

**Title**

The authors may like to consider reworking the title to better reflect the nature of the study... there is nothing new about childbirth satisfaction but there is in modelling the four social psychological determinants they have selected and taking a cross cultural perspective .... Would be worth title reflecting these to some degree. Alternatively you can use the results in the title.

You are probably right. We propose: Assessment of social psychological determinants of satisfaction with childbirth in a cross-national perspective.

---

**Discretionary Revisions (which the author can choose to ignore)**

**Presentation**

Structure

The authors may like to consider reorganising the heading to assist flow of the paper– for example;

- Aim / objectives
- Setting
- Sample size
- Recruitment and data collection
- Measures / Instruments
- Data analysis
- Ethical consideration

We revised the structure of the method section, partly following your suggestion.

Unable to decide on acceptance or rejection until the authors have responded What next?: to the major compulsory revisions

**Level of interest:** An article whose findings are important to those with closely related research interests

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

**Statistical review:** Yes, but I do not feel adequately qualified to assess the statistics.

**Declaration of competing interests:**

I declare that I have no competing interests.
I think this is an interesting and very well conducted study, and addresses a really important topic.

Major Compulsory Revisions (that the author must respond to before a decision on publication can be reached)

My only comments are in relation to the presentation of the analysis.

The presentation of statistical results needs a little attention before it meets modern standards. For example, results should be quoted with a confidence interval. For example, I wish to know whether the confidence intervals for age of the two nationalities of women differs (you quote a mean of 30.4 and 31.9).

The confidence intervals for age do differ for the two nationalities. ACTION: We added confidence intervals to the tables.

The difference between the two in terms of levels of higher education needs looking at very carefully. My first question is whether there is some difference in definitions of higher education?

Indeed, the education systems in Belgium and the Netherlands differ. We took the different definitions of higher education into account when constructing the categories of this variable.

The striking difference in level of education is partly explained by the over sampling of home births in Belgium: most women intending to give birth at home in Belgium are highly educated. ACTION: We added this to the discussion part.

It is the potential for confounding that most interests me throughout this paper.

I see that 68.8 of Belgian women felt the experience met expectations, whereas for Dutch women the figure was 61.1 percent. Again, without confidence intervals I have no way of telling whether or not this is significant, but far more importantly, I am immediately curious as to whether the expectation relates to education levels or nationality, and of course there is strong confounding between nation and education levels.

We agree. We should be more careful with statements based on descriptives only. We added confidence intervals in the tables. We find little support for an association between education levels and expectations in our bivariate statistics (Spearman's rho = 0.066, p-value = 0.114) and cross tabs. We do find some evidence that expectations relate to country in our bivariate statistics (Spearman's rho = -0.10, p-value = 0.014) and cross tabs. This weak association may be caused by the high expectations of Belgian women giving birth at home. But indeed, no conclusions can be drawn until confounding is controled for. ACTION: We added confidence intervals to the tables and we reduced the descriptive part.
When you next tell us that for "45.0% of the primiparae, reality of birth differed from their expectations, compared with 28.7% of the multiparae" I am getting more and more curious about confounding, as there were notable differences (10 percentage points) between first and subsequent deliveries when comparing the two nations.

It is not unusual to find that birth experience differed from expectations more for primiparae than for multiparae (see for example Stolte, 1987), but indeed again no conclusions should be drawn until confounding is controlled for.


We agree. We should not jump to conclusions too quickly. We rewrote the descriptive part, to make sure we do not report findings, which can only be concluded after some variables, such as education, parity, country, are controlled for.

Whilst I appreciate that some descriptive summaries are important at this stage, I feel perhaps they should be tabulated and the comments kept to a minimum, as conclusions are dangerous until you have accounted for all the confounding.

We added the regression equation to the text in the paragraph on data analysis.

We agree. We should not jump to conclusions too quickly. We rewrote the descriptive part, to make sure we do not report findings, which can only be concluded after some variables, such as education, parity, country, are controlled for.

For me, the way to really understand the confounding is by means of the hierarchical models, and that part of your work needs expanding and setting out more carefully to this end.

I really would like to see your model carefully specified, in enough detail for a reader to replicate the analysis with their own data if they so wished. Personally, I prefer to see a formula for the models you are fitting.

We added the regression equation to the text in the paragraph on data analysis.

Whilst you mention that you used SPSS, you give no references for the methodology you have used, or citations to similar work.

Whilst you mention that you used SPSS, you give no references for the methodology you have used, or citations to similar work.

We have revised the paragraph on data analysis and included references to a paper describing the method we used.

Is it appropriate to use the “standard” model, assuming Gaussian responses for the kind of satisfaction data you are attempting to model, or should you use some form of ordinal regression here?

ANOVA is considered to be robust against slight violations of the normality assumption (Chiarotti, 2004). To check whether the violation of the normality assumption had consequences for our analysis, we ran the analysis on transformed dependent variables (Ln transformation of every subdimension of satisfaction). However, comparing the results with our original analysis, no substantial differences occurred. Therefore we decided to present the original analysis, since the interpretation of the results is more straightforward.

As per earlier comments, it is useful (standard?) to report coefficients with confidence intervals, and I would really like to see some reassurance that you have made a number of checks for model fit - it wasn't clear to me how much better models fitted with interaction terms than without, and it wasn't clear which interaction terms you had considered and why.

There are two kinds of interaction terms: first, interaction terms in model 1—this is the model with main effects + interaction term controle*pain, second, interaction terms in model 2—this is model 1 + country specific effects. Within the first model the interaction term controle*pain is added for theoretical reasons. We found in the literature study arguments for including this interaction term (see literature review p. 2 of the manuscript). Within the second model, the country specific effects—this is every determinant*country—are needed to know whether model 1 applies for both countries. This is explained p. 9 under data analysis. The model fit (-2 log restricted likelihood) shows only a very small decrease, which means that the interaction terms do not add much explanatory value to the model (see likelihood ratio test in Table 7 below). We made this explicit to the reader in the description of the results. ACTION: To explain this more clearly we added information to the text under data analysis (p. 9). We also added confidence intervals to the tables.

Minor Essential Revisions (such as missing labels on figures, or the wrong use of a term, which the author can be trusted to correct)

I'm always uncomfortable whenever someone presents data on a number of "correlations". In particular, you are reporting correlations less than 0.3, sometimes a lot less. Does this really tell us much - what do the scatter plots tell you about the strength of the relationships between these variables. I am particularly uncomfortable when these are being estimated in the presence of a number of known confounding factors, and think it makes much more sense to report association in terms of the model coefficients from a model which takes into account confounding and so on.

We agree. We keep the descriptive and bivariate statistics to a minimum and focus on the model coefficients.

Discretionary Revisions (which the author can choose to ignore)

Unable to decide on acceptance or rejection until the authors have responded What next?: to the major compulsory revisions
Level of interest: An article of outstanding merit and interest in its field
Quality of written English: Acceptable
Statistical review: No, the manuscript does not need to be seen by a statistician.
Declaration of competing interests: I have no competing interests.

Version: 1 Date: 18 April 2007
Reviewer: Giorgio Bedogni
Reviewer's report:
General
I was asked to review this manuscript as statistical referee.

Major Compulsory Revisions (that the author must respond to before a decision on publication can be reached)

P2 L7 “Universal character” is too strong. “The applicability of this model to different countries is explored” may be more correct.

OK. We changed this in the text.

P5 L5 Are Ghent and Tilburg comparable for characteristics other than the number of births?

They are comparable on other characteristics such as population density (see tables below) but we thought that the number of births was the most relevant illustration.

P5 L12 Please explain how you performed over sampling of home deliveries.

We contacted all the independent midwives (doing home births) on the list of the VLOV, the Flemish organisation for midwives, by mail and by telephone. First they received an information letter and within two weeks we gave them a call to ask whether they wanted to participate in the study. 59.26% participated (Table 5 below). The expected number of home births in Tilburg in 2004 (N = 720) was approximately the same as the expected number of home births in Flanders (N in 2003 = 704). Therefore we tried to reach as many as possible. We did not use selection criteria: we contacted all midwifery practices throughout Flanders.

P5 L15 What criteria did you use to select midwifery practices outside Ghent?

We contacted all the independent midwives (doing home births) on the list of the VLOV, the Flemish organisation for midwives, by mail and by telephone. First they received an information letter and within two weeks we gave them a call to ask whether they wanted to participate in the study. 59.26% participated (Table 5 below). The expected number of home births in Tilburg in 2004 (N = 720) was approximately the same as the expected number of home births in Flanders (N in 2003 = 704). Therefore we tried to reach as many as possible. We did not use selection criteria: we contacted all midwifery practices throughout Flanders.
If I understand this point correctly, 611 of 833 women who completed an antenatal questionnaire were studied. What was the aim of the antenatal questionnaire for the present study? Did the woman who were followed-up differ from those not followed-up?

One of the aims of our research was to compare expectations before birth with experiences of birth. Therefore we needed an antenatal questionnaire. For the present study the only variable we needed out of the antenatal questionnaire was planned place of birth. We compared the women who dropped out with the women who were followed up, and there were some significant differences between these two groups: More women preferring a hospital birth dropped out than women preferring a home birth. This means that women preferring a home birth were more motivated to participate, or that independent midwives did a better job following up new mothers who hadn't returned the questionnaire yet than the hospital staff. More women who expected painkillers during birth dropped out, but this is related to the hospital context. Women who dropped out had on the average a higher prenatal distress rate, compared to women who were followed up. Since the bivariate correlation between prenatal distress and satisfaction is small (Table 6 below), it can be expected that the women who dropped out would not have been less satisfied than the ones who were followed up.

It is a pity that registration of women who refused to participate was not performed systematically. This makes generalization of the results less likely. The Authors should consider this point in the discussion.

The Mackey Childbirth Satisfaction Rating Scale is made of 6 dimensions that comprise from 2 to 9 items. Each dimension was used as outcome in 6 linear models. Was the outcome variable normally distributed in all cases? This seems unlikely for the dimensions with few items (2 and 3). If this is not the case, a transformation or a model with ordinal outcomes may be tried.

Please, give the number of subjects on whom you performed the reliability study of the scales.
Because you use a mixed model, specify how effects were modelled (what fixed? what random?).
It is very important that the Authors provide a measure of effect size.

OK. We addressed this issue in the discussion.
We tried several transformations to normalize distributions, but the results did not differ from the analysis with untransformed variables. Therefore, we have chosen to report the analysis with untransformed variables. This is more straightforward and easier for the reader to interpret.

The number is 605. We added this to the text.
All determinants are defined as fixed factors. We added this information to the text in the paragraph about the data analysis.
We included the standardized regression coefficient ($\beta$) in table 4 to give an indication of effect size.
What is the variance of each outcome of the Mackey Childbirth Satisfaction Rating Scale explained by covariates?

Although in multilevel designs it is difficult to tell how much variance is explained, we calculated and added R² to table 3 of the manuscript. Hox states: "First, there is unexplained variance at several levels to contend with…. Second, if there are random slopes, the model is inherently more complex, and the concept of explained variance has no unique definition anymore" (Hox, 2002 p. 63). Our model does not include determinants on the country level and no random effects are included, therefore the R² is valid indicator of the variance explained by the covariates. ACTION: we included a small paragraph describing the R² in the results section.


Minor Essential Revisions (such as missing labels on figures, or the wrong use of a term, which the author can be trusted to correct)

P1 L3 “the determinants have been identified”. Isn’t it better to say “some determinants have been proposed?”

OK.

P1 L4 “social psychological”. Is this for “social and psychological determinants”?

No, social psychology is a subdiscipline of sociology. Social psychological determinants are variables related to that domain.

P1 L6 Please explain what you mean by integration in “one model”. This is not clear to me.

We intended to say that these determinants are assessed in one regression equation. ACTION: We reformulated this sentence.

P1 L7 “Located on a single country”. Isn’t it better to write “focuses on single countries”?

OK. We reformulated this sentence.

P2 L2 lower(ed) as compared to what?

The medical model does not result in lower satisfaction scores compared to the social model. ACTION: We reformulated this sentence.

P4 last line: “divergent” needs clarification.

This should be clear now. We revised the background section, where the divergent maternity systems of Belgium and the Netherlands are explained.

P8 L17 “nobody had a place in mind other than the ones summed up”. This is not surprising because the options included “other” and “I don’t know”. Is it intended that these 2 categories were never chosen?

We did not intend to indicate that these categories were never chosen. No one chose "others" and six women chose "I don’t know".
Wouldn’t the analysis benefit from the inclusion of both the planned and actual place of birth?

"No table” means that the result is not presented in one of the tables.

OK. ACTION: We added this sentence to the text: “We present the regression coefficients (B) and the significance of the findings (P-value) between brackets.”

This would lead to multicollinearity problems and does not really add much to the analysis. In another paper (in review), we address the satisfaction with childbirth of women who did not have birth in their place of preference.

"No table” means that the result is not presented in one of the tables.

We added the skewness and kurtosis to the table.

Specify in the text that “B” are regression coefficients.

We added the skewness and kurtosis to the table.

Specify in the text that “B” are regression coefficients.

We added the skewness and kurtosis to the table.

What is meant by “no table” in brackets?

We added the skewness and kurtosis to the table.

We added the skewness and kurtosis to the table.

We omitted the Cronbach's alpha from the tables. They are discussed in the text, but we prefer to keep them integrated with the explanation of each variable, instead of mentioning them separately in a paragraph on reliability.

Table 3 Were these variables normally distributed? If not, mean and SD should not be used to describe them.

Discretionary Revisions (which the author can choose to ignore)

Values of t are not essential and can be removed from the text.

Unable to decide on acceptance or rejection until the authors have responded What next?: to the major compulsory revisions

Level of interest: An article whose findings are important to those with closely related research interests

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

Statistical review: Yes, and I have assessed the statistics in my report.

Declaration of competing interests:

I declare that I have no competing interests