Author’s response to reviews

Title: Mother-Infant Bonding is not Associated with Feeding Type: A Community Study Sample

Authors:
Ilana Hairston (hanahai@telhai.ac.il)
Jonathan Handelzalts (jonathanh@013.net)
Tamar Lehmen-Inbar (tamar.lehman@gmail.com)
Michal Kovo (michalkovo@gmail.com)

Version: 1 Date: 20 May 2018

Author’s response to reviews:

We thank the reviewers for their helpful and insightful comments. We have tried to address each and every comment by answering it directly herein and/or making the required changes in the manuscript. Several of the comments reflect the difficulty with asserting a lack of an association between breastfeeding and bonding as evidence of no association. We would like to emphasize that we acknowledge this difficulty, while we made significant effort to demonstrate that the null hypothesis may – actually - be correct, the potentiality of a Type II error remains. We believe that despite the limitation of the study, these negative findings are as valuable to the scientific community as are positive results.

All changes in the text have been highlighted in red font, however changes in figures and tables have not.

Editor Comments: 1. Method: PBQ: authors need to mention about how the designer of this scale tested the validity and reliability and also, when they translated and back translated, then what did they do with these translations (factor analysis?).


Both citations have now been added to the manuscript.

2. How authors eliminated “incipient abuse“ from questionnaire without any validity test?

This is a good point. Please note that in the original validation paper, by the authors of the scale (Brockington et al., 2006), each subscale was validated separately. Thus, general bonding problems, anxiety about care and rejection subscales, were each found to independently have adequate specificity and sensitivity to the bonding structured interview. Thus, we think that removing the two items, pertaining to abuse, should not substantially impact the validity of the scale.

3. About EPDS and PSQI also authors need to explain how they found these questionnaires are appropriate to use in their country?


Validation of the Hebrew version of the EPDS was done by Glasser, S., Barell, V., 1999. Depression scale for research in and identification of postpartum depression. Harefuah 136, 764 – 768.

These references have been added.

4. Please mention the inclusion/exclusion criteria of study.

Inclusion criteria were willingness to complete the questionnaires in full, exclusion criteria were premature birth (before week 36) and chronic illness of the infant.

This information has been added.
5. Results: table 1: we need to know about the other percentages of education, job and income. For instance how many claimed that their income was sufficient and how many believed that their income was insufficient.

We asked participants how they perceived their household income relative to the national median. We did not ask about perception of sufficiency. More detailed description of education, income and employment status has been added to Table 1.

6. In my opinion, authors should classified all participants into two groups of good bonding and weak bonding and then compare all variables between two groups.

Interesting point. Statistically speaking correlation analysis is used to determine the association between continuous variables, while t-test (or non-parametric equivalents) is used determine the difference between two groups on some variable. It was not the aim of this study to identify the difference between high and low bonding mothers on any of the variables of interest, rather the effects of breastfeeding on bonding. This has now been emphasized in the introduction.

7. Table 3 and 4 are too lengthy and does not permit reader to understand the relation between variables.

We have simplified data presentation doing the following:

a. Added Table 4 depicting the Bayesian Factor regression, to compliment the negative finding regarding lack of correlation between breastfeeding and bonding.

b. Replaced the Table 4, depicting the results of the multivariate analyses on age dependent effects of breastfeeding on PBQ subscales with a figure, and smaller table (Table 5) with the results of the Bayesian ANOVAs only.

c. In the regression table (now Table 6) we removed the p values and added lines between levels of the regression model.

8. Discussion: the first two lines at the beginning of discussion, authors should mention the results of this study with a caution.

We have revised the first paragraph in consideration of this comment.
9. In the limitation of study, “the sample recruited were not representative of breastfed women” should be mention.

While we appreciate this assertion, we disagree with it. Although this was a convenience sample, from higher socio-economic status (which we mention in the limitations), the distribution of breastfeeding behavior was consistent with national reported data.

******

Louise Brough (Reviewer 1): Comments to the authors

This paper describes a cross sectional survey of 271 mothers to investigate the associations between breastfeeding and bonding allowing for negative mood and sleep disturbances.

1. Title
Your measure of bonding was a lack of bonding disorder, rather than a measure of actual bonding. Thus a more appropriate title would be "Mother infant child interaction difficulties are not associated with feeding type."

Changed with some modifications, as we did not measure mother-infant interactions but rather mothers’ emotional bonding experience to their infants.

2. Abstract
Please include information about where the study was carried out in the abstract.

Added

3. Introduction
Breastfeeding rates are country/culture specific; you could include information here about where the data was collected.

Reference to country of data collection has now been added.

4. Methods
The definition of bonding seems to be a lack of bonding disorder, is this truly a measure of bonding.

Indeed, the PBQ is designed to assess bonding disorder. However, Perrelli, Zambaldi, Cantilino, & Sougey (Revista Paulista de Pediatria, 32, 257-265, 2014) recently showed that the PBQ is most often used to assess mother-infant bonding. The second most commonly used instrument
was found to be the MIBS (mother-infant bonding scale), which also assesses bonding difficulties but is only suitable for the first few weeks.

Further, although the PBQ is designed to detect bonding difficulties, several items in the questionnaire refer to positive emotions and behaviors towards the infant, such as “I love my baby very much” or “I enjoy playing with my baby”. Finally, as noted, the report is part of a larger study aimed at assessing the role of several maternal and child characteristics, including depression, sleep problems, etc., and hence the PBQ deemed appropriate for the larger study.

5. How was the sample size determined? There are only 33 women in the never BF group - is this a valid sample size, the lack of findings could be due to insufficient sample size.

Indeed, as noted in the discussion, one of the limitations of the study is the small number of “never breastfed” mothers, even though the proportion was appropriate for the sample size. Despite the small N some effects were found. For example, women who breastfed had more daytime fatigue than women who did not. Further, we used several statistical tools to address the possibility of Type II error, including stratified bootstrap to address differences in groups sizes, and Bayesian statistics. Finally, as noted in the discussion, our results are consistent with others who addressed this question specifically – the relationship between bonding and breastfeeding.

6. 92% were within the normal range - so only 8% had a bonding disorder, thus you would be left with very few in the never group, making your findings less reliable.

Our hypothesis is based on the often-stated belief that breastfeeding improves bonding. We did not hypothesize that lack of breastfeeding is associated with bonding disorder, in fact we have no hypothesis pertaining to bonding disorder.

We were interested in the association of breastfeeding with the degree of affiliative maternal emotions as assessed by the bonding questionnaire, and therefore treated the bonding variable as a continuous variable. Thus, the fact that there are very few women with bonding disorder in the never BF group does not diminish the strength of our conclusion.

Reporting the range and cutoff for bonding disorder was for the sole purpose of generalization of our results, as the PBQ as used in many studies.

7. Why were past BFers not grouped with exclusive and partial - they are all Breastfeeders.

Fair point, although this will hardly change the outcomes. Note, there were no effects of breastfeeding on bonding and no interaction with infant age – where among mothers of younger infants there was a higher proportion of currently breastfeeding mothers. Again, we would like
to highlight the fact that our results contradict a cultural belief, and not the scarce amount of actual data in the field.

8. The sample was a convenience sample. How comparable is your sample to the breastfeeding mothers in Israel? Is your population different/similar in terms of SEC, age, ethnicity to the breastfeeding population in Israel.

As noted in the discussion (now highlighted) the distribution of never, current, past breastfeeding is commensurate with reported national averages, based on National Health and Nutrition Survey from birth to the age of two (2014) from the Israel’s Ministry of Health.

9. Don't use etc - be specific about data collected.

All demographic variables collected are now listed.

10. Is the coupon in US $

The coupon was in Israeli shekels; in the paper the value was translated to US$ to allow the international reader for a better understanding of the value.

11. The statistics used needs to be more clearly explained in the methods section rather than in the results.

More information regarding statistical analyses has been added to the methods section.

12. Results Why was 1 outlier removed can you provide some explanation of why they were deemed an outlier.

As noted in the methods section (now highlighted), we used Chi square distribution of Mahalanobis distance estimates, to identify outliers across measures of interest (i.e., EPDS, total PBQ scores, and sleep symptoms). This analysis identified one outlier.
13. How did those who completed your study compare to those who did not finish the study (585 entries).

Most incomplete entries were due to respondents being screened out when they either reported the age of their infant to be outside the age range of the study, or endorsed exclusionary criteria (chronic illness of the infant or being premature), despite clear description in the ads. Of the remaining respondents who dropped out (approx. 5%), we unfortunately can’t answer this question – as the demographic questionnaire was at the end.

14. Define TST is in the text, I know it is in the table also but need to have in the main section too. The internal consistency for the PSQI was lower (0.583) than for the other 2 measures. Could this have had an appreciable effect on the validity of this instrument?

Definition of TST has been added.

With respect to the internal consistency of the PSQI, our values are consistent with those reported in Shochat, T., Tzchishinsky, O., Oksenberg, A., & Peled, R. (2007), who validated the Israel version of the PSQI. In their report Cronbach-α coefficient in the control group was 0.52.

15. Discussion The not knowing whether the infant is fed expressed BM vs from breast is a major limitation, when you consider the physiology of breastfeeding.

We agree with this assertion and indeed acknowledged it in the limitations section. It should be noted that we aimed to study psychological aspects of breastfeeding rather than the physiology of breastfeeding.

16. You need to discuss whether the sample size is valid. The small number of never BF is a limitation of your study and could be a reason for the lack of association of your findings.

Indeed, as noted in the discussion, one of the limitations of the study is the small number of “never breastfed” mothers, even though the proportion was appropriate for the sample size. Please note that despite the small N, some effects were found. For example, women who breastfed had more daytime fatigue than women who did not. Further, we used several statistical tools to address the possibility of Type II error, including stratified bootstrapping of the data to address differences in groups sizes, and Bayesian statistics. Finally, as noted in the discussion,
our results are consistent with others who addressed this question specifically – the relationship between bonding and breastfeeding.

17. As you have significant limitations you need to be careful with the strength of your assertions about the lack of finding associations between bonding and BF. You need to acknowledge the role breastfeeding plays in nutrition and also other health aspects which are well documented such as the role BM plays in immunology, breast milk is not solely for nutrition.

All the significant limitations are mentioned in the text. Further the role breastmilk plays in the health of the both infant and mother, and in development are mentioned on page 15. We added a reservation regarding these limitations in the closing paragraph of the paper.

Roosy Aulakh (Reviewer 2): Acceptible
Karen Cowgill (Reviewer 3): This article raises an interesting question and highlights a gap in the literature that is important to address - namely, whether breastfeeding promotes mother-infant bonding, as it is commonly believed to. However, the study design and the data presented here are inadequate to answer this question with the degree of confidence the authors assert.

Major concerns: 1. The biggest concern is the self-selection of the study participants, who completed an online survey. Could the authors address this, i.e., how might the self-selection that drove potential respondents to the site or that caused them to participate or not affect the findings (how do mothers of young infants who access and complete a lengthy online survey differ from those who don't)? There are few details about the clinical and social networks through which participants accessed the survey site; just under half of those who entered the survey completed it, suggesting there may be (non-)response bias.

The problem of self-selection in studies and especially in internet studies is a known issue, and shared with most internet-based studies, and indeed we stated this as a limitation of the study. It should be noted however, that the majority of incomplete entries were due to respondents being screened out when either the age of their infant was outside the age range of the study, or endorsing exclusionary criteria (chronic illness of the infant or being premature), despite clear description of inclusion/exclusion criteria in the ads.
Of the remaining respondents who dropped out (approx. 5%), we unfortunately can’t compare them to the rest of the respondents as the demographic questionnaire was at the end.

2. The hypothesis is not clearly specified. It appears that using measures with multiple components enabled the authors to conduct multiple comparisons and pick the component with the highest correlation (Table 3), rather than specifying the associations to test a priori.

The hypotheses are stated in page 6, at the end of the introduction. Each of the hypotheses was tested separately in the results section. We edited the part regarding the first hypothesis for better clarity. Is should be stressed that the correlations between other variables (such as depression, sleep etc’) were not hypothesized a priori, and we used them only to further strengthen our results.

3. If power is the ability to detect an association when it exists, and the post hoc power calculations (Table 4) show such low power for most associations tested (<< 80%), then no conclusions about the associations can be drawn. Failure to detect an association does not prove that the association is absent. The authors acknowledge this (p. 16), but assert their confidence in the results nonetheless. Many of the statistically significant results are for very small effect sizes.

First, please note that the original Table 4 has now been removed and replaced with a figure (Figure 1) and Table 5. Indeed, the absence of an association was the challenge we faced. Arguably, this is a Type II error, due to the many limitations listed by yourself and the other reviewers. However, given that we replicated other observations with respect to bonding (i.e., association with depressive symptoms and sleep symptoms), and implemented a Bayesian approach to hypothesis testing, we believe that our null result reflects a true lack of association.

Our main point is that the link between bonding and breastfeeding, while often asserted, has not been studied sufficiently, unlike other beneficial aspects of breastfeeding. We believe that negative findings are as valuable to the scientific community as are positive results, and of course we welcome attempts to test this association in future, potentially better controlled, studies.

Other concerns:

4. There are several details missing from the methods. What is the larger longitudinal study participants were part of, and where is it described? The translation of the tools is mentioned in
the acknowledgments, but not in the methods - how were translations verified? What did the recruitment ads say? Did all respondents live in Israel?

We have added more detail regarding translation of the instrument in the methods. The larger longitudinal study is yet to be described, as it was only recently completed. However, the aim of the larger study was to follow the development of infant sleep patterns (using parent reports). Breastfeeding was one of the many factors we are following as part of this study.

The ads slightly differed between venues of dissemination, but, basically, they invited parents of healthy full-term infants to participate in a study. There were different ads targeted at ages 1-3 mo, 4-8 mo, and 9-12 mo. When the ads were disseminated via social media, respondent could enter the questionnaires via a link. Otherwise, they were provided an email and phone number of our laboratory.

5. In the Results, more elaboration on the cutoffs and interpretation of the scales used would be helpful. The interpretation of the Bayes Factor does not seem quite accurate; the authors state that "the likelihood of the results given the null hypothesis was 4.5 times greater than for the alternative hypothesis", but isn't this rather the likelihood of the null hypothesis given the data?

Information regarding cutoffs have been added to the methods. Regarding the Bayesian analysis we have rephrased the description accordingly.

6. The authors first (p. 6) contend that infant age acts like a confounder, i.e., that it is related both to the primary exposure (breastfeeding) and the primary outcome (bonding). However, later (p. 11) they describe it as a nuisance variable, which seems less plausible.

A nuisance variable is simply a variable that is correlated with the hypothesized independent variable (i.e., breastfeeding), or in other words – a (type of) confound. JASP software, which was used for Bayesian analyses, does not have a covariate function, thus the only way to control for effects of age was to add it as a nuisance variable. It does not mean that we regard age as a nuisance.

7. Switching between terms is confusing; please use consistent terminology, i.e., "breastfeeding", not "nursing"; "infant", not "baby"; and "human milk", not "breastmilk" (per ILCA-preferred lactation-related language guidelines)
The terms are now in accordance with the ILCA-preferred lactation-related language guidelines.

8. Why was birth order entered as a covariate? What was the hypothesized effect?

Good point. As can be seen in Table 3, birth order weakly correlated with anxiety about care subscale, which was why it was included in this analysis. However, giving that the correlation was weak and potentially spurious, we redid this analysis w/o birth order.

9. The presentation of large tables with raw output makes it difficult to pick out relevant information. Table 2 - why not show bivariate stats, i.e., make columns by breastfeeding status and then show the dependent variables by breastfeeding status? Show total possible scores for PBQ, EPDS, and PSQI. Include numbers as well as %s in all tables; did all participants answer all questions?

Table 2 has been changed accordingly.

10. Table 5 - replace levels 1-3 with never, past, current breastfeeding to facilitate reading.

Currently Table 6. Level 1-3 reflect stages of the hierarchical regression, not the different groups. This has been clarified in the legend.

11. Figure - it looks like 3 individuals with high PBQ scores are drawing the lines up in the current breastfeeding category. Can the authors comment on these, and say whether any outliers were identified/dropped based on their Mahalanobis distance?

The purpose of Mahalanobis distance calculation is to identify multivariate outliers, reflecting unusual combinations of two or more variables in the dataset. In this case, there were indeed three respondents who had higher levels of depression, but this was appropriately associated with higher scores on the bonding scale and more sleep difficulties. Thus, the Mahalanobis distance calculation did not identify these respondents as outliers.
Please note that five mothers had significantly high scores on the PBQ, three of them in the currently breastfeeding group.

12. Several references in the Bibliography are incorrectly formatted or incomplete.

These have been fixed.