

## **The Birth Satisfaction Scale-Revised (BSS-R): Should the sub-scale scores or the total score be used?**

Colin R. Martin<sup>\*1</sup>, Caroline J. Hollins Martin<sup>2</sup>, Ekaterina Burduli<sup>3</sup>,  
Celestina Barbosa-Leiker<sup>4</sup>, Colleen Donovan-Batson<sup>5</sup> and  
Susan E. Fleming<sup>6</sup>

<sup>\*1</sup>Professor of Perinatal Mental Health, Faculty of Health Sciences, University of Hull, Hull, UK, HU6 7RX. E-mail: [C.R.Martin@hull.ac.uk](mailto:C.R.Martin@hull.ac.uk)

<sup>2</sup>Professor in Maternal Health, School of Nursing, Midwifery and Social Care  
Edinburgh Napier University, Scotland, UK. EH11 4BN.  
E-mail: C.HollinsMartin@napier.ac.uk

<sup>3</sup>Post-Doctoral Research Associate, Elson S. Floyd College of Medicine, Washington State University, Spokane, Washington State, USA. E-mail: [eburduli@wsu.edu](mailto:eburduli@wsu.edu)

<sup>4</sup>Associate Professor, Washington State University College of Nursing  
Washington State University, Spokane, Washington State, USA. E-mail:  
[celestina@wsu.edu](mailto:celestina@wsu.edu)

<sup>5</sup>Director, Division of Health Policy and Advocacy, Midwives Alliance of North America, Washington State, USA, E-mail: [Cdonovan-Batson@swtc.edu](mailto:Cdonovan-Batson@swtc.edu)

<sup>6</sup>Assistant Professor, Seattle University College of Nursing  
Seattle, Washington State, USA.  
E-mail: [fleminsu@seattleu.edu](mailto:fleminsu@seattleu.edu)

\*Correspondence

## **The Birth Satisfaction Scale-Revised (BSS-R): Should the sub-scale scores or the total score be used?**

### **Abstract**

**Objective and background:** The 10-item Birth Satisfaction Scale-Revised (BSS-R) is increasingly being used internationally as the instrument of choice for the assessment of birth satisfaction. There remains conjecture over the most appropriate way to score the instrument, the choices being, sub-scale score totals, the overall total score, or adopting both approaches. The current study sought to clarify this issue by examining the measurement characteristics of the United States version of the BSS-R from a large dataset.

**Methods:** A secondary analysis of a data matrix derived from a large sample United States BSS-R validation study (N=2116) using structural equation modelling and bi-factor model evaluation.

**Results:** The bi-factor model revealed an excellent fit to data ( $\chi^2_{(df=25)} = 208.21, p < 0.001, CFI = 0.98, RMSEA = 0.06, SRMR = 0.04$ ), demonstrating the relative independence of the BSS-R quality of care sub-scale, while in contrast the women's attributes and stress experienced during childbearing sub-scales could be explained more plausibly by a general factor of experience of childbirth.

**Conclusion:** Consistent with the recommendations of the original BSS-R validation study, the current investigation found robust empirical evidence to support the use of both the sub-scale total scoring system and the total score. Researchers and clinicians can therefore select either approach (or both) with confidence.

**Key Words:** Birth Satisfaction Scale-Revised (BSS-R), scoring, validation, maternity care, outcomes, midwives, obstetricians

## **Introduction**

Birth satisfaction has been placed centrally within the main domains of interest and relevance for birth outcomes globally (The International Consortium for Health Outcome Measurement, 2017). Given the established relationship of birth satisfaction and key clinical parameters, including delivery type (Fleming et al., 2016), accurate, valid and reliable assessment of the birth satisfaction concept is essential. The Birth Satisfaction Scale-Revised (BSS-R; (Hollins Martin and Martin, 2014) has become established as the 'gold standard' instrument of choice for assessment of birth satisfaction, with the status of the tool conferred by International expert consensus (The International Consortium for Health Outcome Measurement, 2017), validation studies (Barbosa-Leiker, Fleming, Hollins Martin, & Martin, 2015; Fleming, et al., 2016; Goncu Serhatlioglu, Karahan, Hollins Martin, & Martin, 2018; Hollins Martin and Martin, 2014; Jefford, Hollins Martin, & Martin, 2018; Martin et al., 2017) and clinical investigations highlighting the veracity and applicability of the measure (Hinic, 2016, 2017).

A topical issue since the original development of the BSS-R (Hollins Martin and Martin, 2014) is which scoring approach to undertake. The BSS-R assesses three domains of (i) *stress experienced during childbearing*, (ii), *women's attributes*, and (iii), *quality of care*. Hollins Martin and Martin (2014) recommended in their original paper that sub-scale scores representing the above domains can be used and the developers offer a scoring algorithm to facilitate this. Additionally, Hollins Martin and Martin (2014) also suggest that the total 10-item score is permissible as an overall index of birth satisfaction or the experience of childbearing. Consequently, those using the BSS-R are left with essentially three choices for scoring, (i) sub-scales, (ii) total score or (iii) sub-scales and total score.

It was noted that though the original study established a three-factor correlated model as the measurement model of the tool, thus the rubric for the three derived sub-scales, a second-order model was also evaluated with experience of childbearing being the second-order factor (Hollins Martin and Martin, 2014). Importantly, support for a second-order model would suggest that the variance in the three domains would be explained by the second-order factor and thus would imply a (statistical) preference for the use of a total score. However, no evidence was found for the superiority of this higher-order model when compared to the three-factor correlated model which remains the foundation for the established multi-dimensional scoring system of the BSS-R (Hollins Martin and Martin, 2014).

Given that the BSS-R is being increasingly used world-wide, is globally endorsed for clinical outcome assessment (The International Consortium for Health Outcome Measurement, 2017) addressing the scoring preferences of the BSS-R empirically is both prudent and relevant.

The lack of a definitive preference for a model type may be accounted for by the inherent limitations of the higher-order model and its specification (Chen, West, & Sousa, 2006), thus the observation of little difference between models in model-fit when compared. An alternative approach has been suggested, the bi-factor model, which offers a number of advantages over higher-order models and to some degree correlated models (Chen, et al., 2006). The bi-factor model tests fit to data by examining how much of the variance in all of the items of the scale are explained by a general factor, and within the model specification, how much of the *remaining* variance in the items is explained by domain-specific factors (Chen, et al., 2006).

The current investigation sought to generate an empirically-driven recommendation regarding the scoring of the tool. The objectives of the current study were to:

1. Compare the three-factor correlated model of the BSS-R with an alternative bi-factor model.
2. Determine whether a general factor of 'experience of childbearing' explains the variance in the items of the BSS-R.
3. Determine whether domain-specific factors of 'stress experienced during childbearing', 'women's attributes' and 'quality of care' explains the variance in the items of the BSS-R.
4. Based on the findings from objectives 1-3, recommend an empirically-derived scoring approach to the BSS-R.

## Method

A secondary analysis of the data matrix from the United States large sample BSS-R study (Fleming, et al., 2016; Martin, et al., 2017).

## Participants

Participant characteristics (N = 2116) are described in detail in (Fleming, et al., 2016; Martin, et al., 2017), with pertinent details of study design, demographics, clinical details, and ethical approval summarised.

## Measures

The BSS-R (Hollins Martin and Martin, 2014) is scored on a five-point Likert type scale with responses ranging from (i.) strongly agree, (ii.) agree, (iii.) neither agree or disagree, (iv.) disagree, (v.) strongly disagree, with reverse scoring of four items. The three BSS-R sub-scales include *stress experienced during childbearing* (four items) e.g. 'I found giving birth a distressing experience', *quality of care* (four items), e.g. 'The delivery room was clean and hygienic' and *women's attributes* (two items), e.g. 'I felt out of control during my birth experience'.

## **Statistical analysis**

The objectives of the study were addressed using a structural equation modelling (SEM) approach to model evaluation (Byrne, 2010), specifically confirmatory factor analysis (CFA). The tri-dimensional measurement model of the BSS-R comprising correlated factors of *stress experienced during childbearing*, *women's attributes* and *quality of care* specified by Hollins Martin and Martin (2014) was compared to an alternative bi-factor model comprising a *general* factor and the three domain-specific factors previously outlined. Model fit indices (Bentler and Bonett, 1980) used to evaluate fit included the comparative fit index (CFI; Bentler, 1990), the root mean squared error of approximation (RMSEA; (Steiger and Lind, 1980) and the square root mean residual (SRMR; (Hu and Bentler, 1999). The relationship between both domain-specific factors and the general factor within the bi-factor model are specified as orthogonal, since an essential characteristic of the model is to identify the additional contribution of the domain-specific factor beyond that of the general factor.

### **Preparation of the data covariance matrix**

A correlation matrix from the United States large sample BSS-R dataset of the 10 BSS-R items was obtained from the instrument developers (Hollins Martin and Martin, 2014). The standard deviations from each item of the BSS-R were also obtained. Knowledge of the standard deviations of each of the BSS-R items allows the correlation matrix to be converted into a covariance matrix. Information on the sample size from the United States large sample BSS-R dataset (N=2116) is then incorporated into the data analysis with the covariance matrix and this provides an accurate approximation of the original dataset for the purposes of SEM without the need to access the original dataset itself.



## Results

The three-factor measurement model of the BSS-R offered a good fit to the data ( $\chi^2_{(df=32)} = 256.82$ ,  $p < 0.001$ , CFI = 0.98, RMSEA = 0.06, SRMR = 0.04)<sup>1</sup>. The bi-factor model (Figure 1.) revealed fit statistics similar to those of the three-factor correlated model ( $\chi^2_{(df=25)} = 208.21$ ,  $p < 0.001$ , CFI = 0.98, RMSEA = 0.06, SRMR = 0.04). Examination of item-factor loadings revealed that the quality of care factor satisfied the criteria for being a definable independent factor (item-factor loadings  $> 0.3$ ). However, the women's attributes and stress experienced during childbearing factors were more convincingly explained by the general factor (item-factor loadings  $> 0.3$ ). The general factor was observed to explain a significant proportion of the variance in BSS-R items 1., 2., 7., 9., 4., 8., 3., 5., and 6.

---

<sup>1</sup>As would be anticipated, model-fit of the three-factor correlated model is identical to that reported by Martin et al. (2017), the trivial fractional variation in  $\chi^2$  (256.82 vs. 255.15) explained by rounding to two decimal points in the correlation matrix and individual BSS-R item standard deviations.

A posteriori single-factor CFA was therefore conducted to examine the model-fit of these nine items. This revealed a poor fit to data ( $\chi^2_{(df=27)} = 2546.51, p < 0.001, CFI = 0.72, RMSEA = 0.21, SRMR = 0.10$ ). A further posteriori single-factor CFA was conducted comprising stress experienced during childbearing and women's attributes items with excellent fit to data being observed ( $\chi^2_{(df=9)} = 77.15, p < 0.001, CFI = 0.99, RMSEA = 0.06, SRMR = 0.02$ ). The quality of care items were evaluated within a single-factor CFA ( $\chi^2_{(df=2)} = 12.35, p = 0.002, CFI = 0.99, RMSEA = 0.05, SRMR = 0.01$ ) and revealed excellent fit to data. A two-factor correlated model comprising quality of care items and a combined factor of women's attributes and stress experienced during childbearing was evaluated with excellent fit to data being observed ( $\chi^2_{(df=34)} = 272.81, p < 0.001, CFI = 0.99, RMSEA = 0.06, SRMR = 0.04$ ). Finally, comparison of the three-factor correlated measurement model of the BSS-R and the two-factor model outlined above was conducted using the  $\chi^2$  differences test revealing the three-factor correlated model to provide a significantly better fit to data ( $\Delta\chi^2 = 15.99_{(\Delta df=2)}, p < 0.001$ ).

FIGURE 1. ABOUT HERE

## Discussion

The findings from this reanalysis of the United States large sample BSS-R dataset offers some useful insights into scoring approaches that can be taken with the BSS-R. The bi-factor model was shown to offer a similar fit to data in comparison to the three-factor correlated model that circumscribes the measurement model of the BSS-R. It was noted that the variance of the two items comprising the women's attributes factor were explained by the general factor, with the domain-specific factor of women's attributes contributing little explanatory variance to BSS-R items 4 and 8 once the influence of the general factor is taken into account. A similar observation was observed for the stress experienced during childbearing sub-scale. This contrasts with the quality of care sub-scale which demonstrated independence as a domain-specific factor once the impact of the general factor was taken into account. The findings from the current study indicate the value of the BSS-R domain-specific sub-scales as valuable and independent measures of discrete aspects of birth satisfaction, particularly in relation to the quality of care sub-scale. The finding that the general factor accounted for a significant proportion of the variance in nine of the BSS-R items would also indicate the valid use of the BSS-R total score. The contribution of the general factor to the variance of most BSS-R items offers a clear indication that the overall score measures a coherent theme of satisfaction with the experience of childbearing, as is suggested by Hollins Martin and Martin (2014) in their original instrument development study and their evaluation of a higher-order model that considers this notion. The relative independence of the quality of care sub-scale also provides an indication that this four-item sub-scale could be used separately to the BSS-R, should a robust, short and valid self-report measure of quality of care be required, thus a standalone self-report measure of care quality.

Limitations of the study included using a covariance matrix to approximate actual data, however in the very specific context of SEM, this approximation is extremely accurate to actual data and fit for this specific purpose. A strength of the study is the additional insights into the BSS-R gained using bi-factor modelling, an approach which has rarely been used in reproductive psychology yet offers some valuable potential in understanding the measurement characteristics of tools used in the area.

## References

- Barbosa-Leiker, C., Fleming, S., Hollins Martin, C. J., & Martin, C. R. (2015). Psychometric properties of the Birth Satisfaction Scale-Revised (BSS-R) for US mothers. *Journal of Reproductive and Infant Psychology*, 33(5), pp. 504-511. doi:10.1080/02646838.2015.1024211 Retrieved from <http://dx.doi.org/10.1080/02646838.2015.1024211>
- Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, 107(2), pp. 238-246. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/2320703>
- Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the evaluation of covariance structures. *Psychological Bulletin*, 88, pp. 588-606.
- Byrne, B. M. (2010). *Structural Equation Modeling with AMOS: Basic Concepts, Applications and Programming* (2nd ed.) New York: Routledge/Taylor and Francis Group.
- Chen, F. F., West, S. G., & Sousa, K. H. (2006). A comparison of bifactor and second-order models of quality of life. *Multivariate Behavioral Research*, 41(2), pp. 189-225.
- Fleming, S. E., Donovan-Batson, C., Burduli, E., Barbosa-Leiker, C., Hollins Martin, C. J., & Martin, C. R. (2016). Birth Satisfaction Scale/Birth Satisfaction Scale-Revised (BSS/BSS-R): A large scale United States planned home birth and birth centre survey. *Midwifery*, 41, pp. 9-15. doi:10.1016/j.midw.2016.07.008 Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/27494569>

- Goncu Serhatlioglu, S., Karahan, N., Hollins Martin, C. J., & Martin, C. R. (2018). Construct and content validity of the Turkish Birth Satisfaction Scale - Revised (T-BSS-R). *Journal of Reproductive and Infant Psychology*, pp. 1-11. doi:10.1080/02646838.2018.1443322 Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/29553295>
- Hinic, K. (2016). Predictors of Breastfeeding Confidence in the Early Postpartum Period. *Journal of Obstetric, Gynecologic, and Neonatal Nursing*, 45(5), pp. 649-660. doi:10.1016/j.jogn.2016.04.010 Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/27472996>
- Hinic, K. (2017). Understanding and Promoting Birth Satisfaction in New Mothers. *MCN: American Journal of Maternal Child Nursing*, 42(4), pp. 210-215. doi:10.1097/NMC.0000000000000345 Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/28654447>
- Hollins Martin, C. J., & Martin, C. R. (2014). Development and psychometric properties of the Birth Satisfaction Scale-Revised (BSS-R). *Midwifery*, 30(6), pp. 610-619. doi:10.1016/j.midw.2013.10.006 Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/24252712>
- Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6, pp. 1-55.
- Jefford, E., Hollins Martin, C. J., & Martin, C. R. (2018). Development and validation of the Australian version of the Birth Satisfaction Scale-Revised (BSS-R). *Journal of Reproductive and Infant Psychology*, 36(1), pp. 42-58. doi:10.1080/02646838.2017.1396302 Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/29517299>

Martin, C. R., Hollins Martin, C. J., Burduli, E., Barbosa-Leiker, C., Donovan-Batson, C., & Fleming, S. E. (2017). Measurement and structural invariance of the US version of the Birth Satisfaction Scale-Revised (BSS-R) in a large sample. *Women and Birth, 30*(4), pp. e172-e178. doi:10.1016/j.wombi.2016.11.006  
Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/27965174>

Steiger, J. H., & Lind, J. (1980) *Statistically-based tests for the number of common factors*. Paper presented at the Annual Spring Meeting of the Psychometric Society, Iowa City, USA.

The International Consortium for Health Outcome Measurement. (2017). *Pregnancy and Childbirth Standard Set and Reference Guide, version 1.0.3*.  
<http://www.ichom.org/medical-conditions/pregnancy-and-childbirth/>