

**Translation and validation of the Dutch version of Birth Satisfaction Scale-
Revised (BSS-R)**

Berbel Emmens¹

Caroline J. Hollins Martin²

and

Colin R. Martin*³

¹Independent Researcher, counselor, MSc Applied Psychology

²Professor in Maternal Health, School of Nursing, Midwifery and Social Care
Edinburgh Napier University, Scotland, UK. EH11 4BN.

³Professor of Perinatal Mental Health, Institute for Clinical and Applied Health
Research (ICAHR), University of Hull, UK, HU6 7RX.

***Corresponding author**

Address for correspondence:

Professor Colin R. Martin,

Institute for Clinical and Applied Health Research (ICAHR)

Rm 329, Allam Medical Building

University of Hull

Hull, HU6 7RX, UK

Email: C.R.Martin@hull.ac.uk

Keywords: Birth satisfaction, birth experience, scales, psychometrics, trauma

Abstract

Background: Satisfaction with the birth experience has been recognized to be of significant importance to the well-being of both the mother and her relationship with the baby. Recent observations have also suggested birth satisfaction may be significantly associated with postpartum post-traumatic stress disorder (PP-PTSD). The Birth Satisfaction Scale-Revised (BSS-R) is increasingly used Internationally as a short, valid and reliable multi-dimensional measure of birth experience. The current study sought to develop a Dutch version of the BSS-R (D-BSS-R) for clinical and research application in the Netherlands.

Methods: Post-translation a cross-sectional design with an embedded between-subjects component was used to evaluate key indices of validity and reliability of the D-BSS-R in a purposive sampled cohort of 244 Dutch-speaking women in the Netherlands. Confirmatory factor analysis, divergent, convergent and known-groups discriminant validity were evaluated as was the internal consistency of the measure.

Results: The D-BSS-R was found to be a generally valid and reliable measure of birth experience with the key measurement characteristics of the original English-language measure transferring well to the Dutch context. **Statistically** significant negative correlations were observed between all D-BSS-R sub-scales and a validated measure of PTSD.

Conclusions: The D-BSS-R represents a valid and reliable measure of birth experience suitable and appropriate for use in the Netherlands. The study corroborates previous suggestions of linkage between birth satisfaction and PP-PTSD using a robust and diagnostically valid measure of trauma. The direction of future research is indicated.

Introduction

The experience of childbirth is consistently described as an event of great psychological importance in a woman's life (Hoffman & Banse, 2020). However, potential negative aspects of birth experience can include post-partum post-traumatic stress disorder (PP-PTSD) with deleterious impact for both the mother herself and the quality of her relationship with her child (Williams, Patricia Taylor, & Schwannauer, 2016). PP-PTSD can cause changes in the woman's physical well-being, mood, behavior and social interaction, breastfeeding and have a negative influence on the relationship with the partner and the desire to have children (Ayers, Bond, Bertullies, & Wijma, 2016).

Approximately a third of women rate their delivery as psychologically traumatic (Ayers & Pickering, 2001; Boorman, Devilly, Gamble, Creedy, & Fenwick, 2014). It has long been recognised that care providers may not recognise symptoms of psychological and emotional trauma because of their perception that birth trauma is a physical injury (Beck, 2004a, 2004b). A recent large study (N=2192) of Dutch mothers who had experienced a traumatic birth experience identified lack or loss of control and interaction with the caregiver (regarding communication and emotional and practical support) as the main cause of their traumatic birth experience indicated (Hollander et al., 2017). Satisfaction is an important clinical outcome and is one of the most commonly reported outcome measures for quality of care and health care provided (Sawyer et al., 2013). Birth satisfaction has long been recognised as a key indicator of the quality of maternity care (Hodnett, 2002). Birth satisfaction represents a sophisticated multi-dimensional construct comprising discrete but related aspects of perceived stress related to the birth experience, innate characteristics of the

woman herself and appraisal of the quality of care received (Hollins Martin & Martin, 2014). A recent study in Israel found a significant relationship between births perceived as traumatic and birth satisfaction (Skvirsky, Taubman-Ben-Ari, Hollins Martin, & Martin, 2019) and large National study in the UK found that both post-partum post-traumatic stress and general post-traumatic stress was significantly associated with birth satisfaction (Harrison, Ayers, Quigley, Stein, & Alderdice, 2020). Hollander et al.'s (2017) study highlighted the importance of adequate communication and support of the caregiver to prevent PP-PTSD within the Dutch context and its relevance to the maternal mental health evidence base, these findings being corroborated by the observations of Skvirsky et al. (2019) and Harrison et al. (2020). The Israeli and UK studies used the Birth Satisfaction Scale-Revised (BSS-R)(Hollins Martin & Martin, 2014), a short, valid and reliable multi-dimensional measure of birth experience. The BSS-R is recommended as the self-report measure of choice for measuring birth satisfaction by the *International Consortium of Health Outcomes Measurement* (ICHOM), within the *Pregnancy and Childbirth Standard Set* (The International Consortium for Health Outcome Measurement, 2016). The BSS-R has been widely translated and validated internationally for example, the United States (Barbosa-Leiker, Fleming, Hollins Martin, & Martin, 2015), Turkey (Goncu Serhatlioglu, Karahan, Hollins Martin, & Martin, 2018), Greece (Vardavaki, Hollins Martin, & Martin, 2015), Italy (Nespoli et al., 2020), Spain (Romero-Gonzalez et al., 2019), Slovakia (Skodova, Nepelova, Grendar, & Baskova, 2019), Australia (Jefford, Hollins Martin, & Martin, 2018), Iran (Mortazavi, Mehrabadi, Hollins Martin, & Martin, 2020) and Pakistan (Zafar et al., 2021), however to date, a validated Dutch version of the measure has not been available. The purpose of the current study is thus two-fold, firstly to develop and

validate BSS-R into Dutch and secondly to examine the relationship between birth satisfaction and symptoms associated with PP-PTSD.

To validate the Dutch BSS-R (D-BSS-R) the study had the following objectives:

1. Determine the transferability of the tri-dimensional measurement model of the BSS-R to the D-BSS-R.
2. Evaluate the internal consistency of D-BSS-R Quality of Care (QC), Women's Attributes (WA), and Stress Experienced during Childbearing (SE) sub-scales and the total D-BSS-R scale.
3. Determine the known-groups discriminant validity of the D-BSS-R adopting the same approach used in the original UK BSS-R development study.
4. Determine the divergent validity of the D-BSS-R utilising the approach taken in the original BSS-R development study.
5. Evaluate the relationship of the D-BSS-R to post-traumatic stress postpartum.

Method

Design

The study used a cross-sectional study design to address objectives 1-4 and a subset of this cohort to evaluate objective 5. Inclusion criteria included having given birth within the past 5 years, this based on a review paper indicating reliability and validity of the BSS-R at up to 5 years postpartum (Alfaro Blazquez, Corchon, & Ferrer Ferrandiz, 2017). In the Netherlands, the maternity care system is divided into three levels of care. Healthy women with a low-risk pregnancy receive care from independent community midwives or general practitioners during pregnancy and childbirth (primary care), while high-risk women (or those who become high-risk during pregnancy or childbirth) receive care from an obstetrician in a hospital setting

(secondary care). The tertiary care contains obstetricians and clinical midwives in academic hospitals.

Participants

Participants were eligible for inclusion in the study if they were 18 years or older, able to complete Dutch questionnaires and if their delivery occurred at least one month and at most 5 years ago.

Ethical approval

Ethical approval was requested by the Medical Ethics Review Committee [REDACTED]. The METC [REDACTED] confirmed exemption from the Medical Research Involving Human Subjects Act [REDACTED].

Measures

To measure birth satisfaction, participants were asked to complete the Birth Satisfaction Scale-Revised (Hollins Martin & Martin, 2014). This questionnaire was specifically chosen because the 10-item Birth Satisfaction Scale-Revised (BSS-R) is a widely used, valid and reliable birth satisfaction questionnaire. The BSS-R is recommended by international experts worldwide as an outcome measure for birth satisfaction (ICHOM, 2016). This BSS-R was translated from English into Dutch by two independent translators, the researcher and an expert in maternity care. The Dutch version was backwards translated by one independent bilingual translator whose mother tongue is English. The discrepancies found after the backward translation were discussed and advice was sought from the developers of the BSS-R (Hollins Martin & Martin, 2014). Consensus was reached on the final version of the translation. The questionnaire has three different but related subscales; quality of

care provision (4 items, QC sub-scale), women's personal attributes (2 items, WA sub-scale), and stress experienced during labour (4 items, SE sub-scale).

To measure the severity score on birth trauma, participants were asked to complete the PCL-5 questionnaire (Weathers et al., 2013). The PCL-5 is a 20-item self-report questionnaire that measures the 20 symptoms of PTSD according to the DSM-5 (APA, 2013). The PCL-5 is a widely used valid and reliable questionnaire to measure 20 symptoms of PTSD according to the DSM-5 (APA, 2013), without asking about criterion A, the traumatic event. The Dutch translation of the PCL-5 (Boeschoten, Bakker, Jongedijk, & Olf, 2014) was used in the current study, where the stressful event was related to the delivery.

Procedure

Participants were recruited via several birth-related social media accounts and midwifery practices within the Netherlands and were asked to complete two questionnaires. The first questionnaire (BSS-R) was preceded by some demographic questions. The second questionnaire (PCL-5) was preceded by a detailed informed consent as these questions could be potentially disconcerting in case of a negative childbirth experience. **Questionnaire completion took between 10-15 minutes.**

Statistical analysis

Confirmatory Factor Analysis

Objective 1 was evaluated using confirmatory factor analysis (CFA). CFA is a parametric technique and thus evaluation of data distributional characteristics is important to determine normality and suitability for this approach (Brown, 2015). Individual BSS-R items are thus screened for excessive skew and kurtosis and multivariate outliers are identified and removed (P. Kline, 2000). Previous validation studies of the BSS-R have identified very few outliers within datasets and individual items free from excessive skew and kurtosis (Jefford et al., 2018; Nespoli et al., 2020; Romero-Gonzalez et al., 2019). The tri-dimensional measurement model of the BSS-R (Hollins Martin & Martin, 2014) comprising correlated factors of SE, WA and QC was evaluated by CFA as was a bi-factor model comprising uncorrelated SE, WA and QC and a general factor of birth experience. Recent observations of a good fit to data of the bi-factor model have provided additionally measurement evidence to support the utility of the BSS-R sub-scales and the suitability for using the total BSS-R score (Martin et al., 2018). A single-factor model was also evaluated. Maximum-likelihood estimation (Brown, 2015; R. B. Kline, 2011) was used to evaluate models and model fit adequacy determined using the comparative fit index (CFI) (Bentler, 1990), the root mean squared error of approximation (RMSEA)(Steiger & Lind, 1980), and the square root mean residual (SRMR) (Hu & Bentler, 1999). Cut-off values of >0.90 (CFI), <0.08 (RMSEA) and <0.06 (SRMR) were used as reference threshold values for model acceptability.

Internal consistency

Internal consistency of the D-BSS-R sub-scales SE and QC, and total score was Cronbach's alpha (Cronbach, 1951) was used to evaluate the internal consistency of the SE and QC sub-scales and the whole D-BSS-R scale. A conventional threshold of 0.70 or greater is deemed acceptable (P. Kline, 2000). Comprising two items the inter-item correlation (Pearson's r) with threshold values of 0.15-0.50 was used to determine internal reliability of the WA sub-scale (Clark & Watson, 1995). To facilitate comparison with Hollins Martin and Martin's (2014) original study, Cronbach's alpha was also calculated for the WA sub-scale.

Known-groups discriminant validity

Several previous validation studies (Fleming et al., 2016; Romero-Gonzalez et al., 2019; Skvirsky et al., 2019; Vardavaki et al., 2015) have evaluated known-groups discriminant validity (KGDV) of the BSS-R by comparison of BSS-R scores between groups dichotomized on the basis of delivery type, thus spontaneous vertex delivery (SVD), in comparison to intervention delivery (ID; elective Caesarean section (CS), emergency CS, suction cap and instrument). This approach was adopted for the current investigation. Additionally, comparison between Caesarean section type (elective vs. emergency) was undertaken since potential differences have been occluded in the dichotomous categorization approach of previous studies.

Divergent validity

To determine divergent validity D-BSS-R sub-scale and sub-scale scores were correlated (Pearson's r) with participant age. No statistically significant correlations were anticipated.

Convergent validity

To determine divergent validity D-BSS-R sub-scale and sub-scale scores were correlated (Pearson's r) with the PCL-5 checklist for DSM criteria PTSD. Statistically significant *negative* correlations were predicted between D-BSS-R sub-scale and total scores and the PCL-5 total score.

Results

Participants

Two-hundred and forty-four women consented to take part in the study. D-BSS-R multivariate outliers were identified by reference to Mahalanobis distances (N=3) and removed from the dataset, thus the dataset for psychometric analysis comprised N=241 participants (mean age was 31.94 (SD 4.16), gestational age 39.41 (SD 2.05) weeks). The descriptive and distributional characteristics of D-BSS-R items, sub-scales and total scores are summarized in Table 1. and reveal no evidence of excessive skew or kurtosis. Complete PCL-5 data was provided by N=127 participants.

TABLE 1. ABOUT HERE

Confirmatory factor analysis

The single-factor model (model 1.) revealed a poor-fit to data. The three-factor model also revealed a modest fit to data across indices with acceptable CFI and SRMR, but sub-optimal RMSEA (model 2.). Examination of modification indices did not reveal approaches to improving model fit which were theoretically cogent with the BSS-R measurement model. All items loaded significantly onto their respective factor with the exception of item-10 'The delivery room was clean and hygienic' (Figure 1.). It was observed that SE and WA factors were highly correlated (~1.00) and consequently a posteriori two-factor model with combined SE and WA items as a

single factor and QC items as a second factor was run, $\chi^2_{(df=34)} = 116.30$, RMSEA = 0.100, SRMR = 0.073, CFI = 0.928. The χ^2 differences test revealed no statistically significant differences between the two-factor and three-factor models ($\Delta\chi^2_{(df=2)} = 3.51, p = 0.17$). In contrast, the bi-factor model revealed an excellent fit to data in terms of CFI and SRMR indices while RMSEA was found to be borderline acceptable. Scrutiny of item-factor loadings indicated a strong general factor of combined SE and WA items and an independent QC factor.

FIGURE 1. ABOUT HERE

TABLE 2. ABOUT HERE

D-BSS-R sub-scale and total score correlations

All correlation combinations were statistically significant ($p < 0.01$). Using the correlational comparison method of Diedenhofen and Musch (2015) revealed statistically significant differences between the current study and the original UK BSS-R development study, with the exception of SE-WA ($p = 0.36$) (Table 3.).

TABLE 3. ABOUT HERE

Internal consistency

Cronbach's alpha of D-BSS-R total scale and all sub-scales were all observed to exceed threshold (> 0.70). No significant difference was observed between internal consistency estimations of the current study and those of Hollins Martin and Martin (2014) with the exception of the total scale where alpha was significantly greater (Table 4.). Inter-item correlation of the D-BSS-R sub-scale WA items was $r = 0.57$, $p < 0.001$.

TABLE 4. ABOUT HERE

Known-groups discriminant validity

The spontaneous vertex delivery group were observed to have statistically significant BSS-R scores across all sub-scales and total score, compared to those in the intervention group. Effect sizes were large, with the exception of the QC sub-scale (medium) (Table 5.).

TABLE 5. ABOUT HERE

Significantly higher scores were observed on SE and QC sub-scales and the total D-BSS-R score in the elective Cesarean section group compared to the emergency Cesarean section group, and effect sizes were observed to be large (Table 6.). No statistically significant difference was observed between groups on the WA sub-scale.

TABLE 6. ABOUT HERE

Divergent validity

No significant correlations were observed between WA and QC sub-scales and the D-BSS-R total score and participant age (WA $r = 0.06$, $p = 0.36$, QC $r = 0.006$, $p = 0.92$, and Total $r = 0.08$, $p = 0.20$). However, age and the SE sub-scale score were observed to be significantly correlated, $r = 0.13$, $p = 0.05$, thus with increasing age participants reported less stress.

Convergent validity

Correlations between D-BSS-R total and sub-scale scores and the PCL-5 total score were all observed to be statistically significant ($p < 0.001$) and negative (Table 7.).

TABLE 7. ABOUT HERE

Post-hoc analysis¹

A linear regression was undertaken with the D-BSS-R total score and time since delivery entered as predictors and PCL-5 score as the dependent variable. The model was found to be statistically significant, $R^2 = .26$, $F(2,124) = 21.74$, $p < .001$.

¹ This analysis was undertaken at the suggestion of a reviewer of the original manuscript and thus is post-hoc to the original statistical analysis plan. We are grateful to the reviewer for the suggestion and the additional insights this analysis yielded.

The D-BSS-R total score was observed to be a significant predictor of the PCL-5 score, $b = -0.84$, 95% CI $[-1.10, -0.59]$, $t(124) = -6.58$, $p < .001$, $pr^2 = .26$. In contrast, time since delivery did not significantly predict PCL-5 score, $b = -0.01$, 95% CI $[-0.02, 0.00]$, $t(124) = -1.20$, $p = .232$, $pr^2 = .01$.

Discussion

Results from this study indicate that the D-BSS-R is a psychometrically robust tool for maternity care professionals to evaluate women's experiences of childbirth in the Netherlands. The statistical results suggest that the 3-factor model generally offers a somewhat ambiguous fit to the data which based on two of the three fit indices, would suggest that in addition to total score, the three subscales of the D-BSS-R can be used independently. However, we are minded that the RMSEA values for the three-factor measurement model and bifactor model suggest a caveat for drawing a conclusion of unambiguous good fit to data. To view associated questions please refer to Table 1. The same caveat we would also draw in relation to the bi-factor model findings in terms of using the D-BSS-R total score (Martin et al., 2018). Hu and Bentler (1999) advocate that model fit veracity should be decided on balance across fit indices in instances of conflicting findings. SRMR and CFI were both acceptable for the three-factor and bi-factor models and therefore, with a modicum of caution highlighted, we would conclude on balance that model fit is acceptable for these models, though further research is required to address this conclusion with greater confidence. Many previous studies of the psychometric properties of the BSS-R have generally found good fit to the three-factor measurement model across all model fit indices, therefore we would suggest further investigation of the measurement model characteristics of the Dutch version in new populations. A curious finding was the observation that item 10. *'The delivery room was clean and hygienic'*, did not load onto the QC factor. Interestingly, this item had the highest mean score, smallest SD and range, and greatest skew and kurtosis of all of the items with 70% of participants endorsing the highest item score and less than 1% the lowest. It may be that this finding is specific to the Dutch context of birth where

expectations of a high-quality birthing environment are generally met by high quality service provision (van Stenus, Boere-Boonekamp, Kerkhof, & Need, 2017). Further these item 10 distributional characteristics *per se*, may contribute statistically to the absence of loading onto the QC factor. The finding of a high correlation between SE and WA sub-scales highlights the possibility that the D-BSS-R might best be conceptualised within a two-factor measurement model. However, no statistical differences between three-factor and two-factor models were observed and moreover, a high correlation between these two factors was observed in the original validation study (Hollins Martin & Martin, 2014) and these observations should be contextualised within both the theoretical framework of the BSS-R and the independent utility of the factor-derived sub-scales. It should be noted also that other studies, for example Nasiri, Kariman, and Ozgoli (2020) have shown a much lower correlation between SE and WA sub-scales. Interestingly, Martin et al. (2018) evaluated this two-factor correlated model based on their observations from their bi-factor model analysis and found an excellent fit to data, however in that instance the three-factor correlated model was a statistically significant better fit. Absolute comparison of fit indices suggests the bi-factor model offers a better fit to data compared to the three-factor model. However, an important rider to drawing a conclusion based on these comparisons has been highlighted by Murray (2013) in terms of an inherent statistical bias in favour (in terms of fit indices) of bi-factor models over non-bi-factor measurement models. Based on extensive model evaluation using simulations Murray (2013) cautions against determination of the most appropriate model based on model fit indices alone in these circumstances. This position is also supported by the observation of the high correlation between SE and WA factors which within a bi-factor modeling framework would support the

notion of these combined SE and WA items being indicative of a general factor, thus supporting that aspect of a bi-factor model. The QC sub-scale items within the bi-factor model clearly represent a distinct factor from the general factor thus the notion that a bi-factor model represents a more appropriate interpretive account would not be supported on the basis of a general factor explaining the whole scale within a uni-dimensional context, since there is clear evidence of multi-dimensionality within the measure. Indeed, our findings are consistent with those of the bi-factor model of the US version of the BSS-R undertaken by Martin and colleagues (2018) in terms of the alignment to a general factor of SE and WA items and a separate QC factor. Further work is required to consider the implications of the bi-factor model in terms of any future revision of the D-BSS-R and indeed the BSS-R more widely, a germane issue to that future discourse being the degree of correlation found between SE and WA sub-scales across different language versions and contexts where the BSS-R is used. The modest RMSEA and some evidence of a ceiling effect in relation to item 10. may be of interest to others undertaking translation/validation studies of the BSS-R in terms of estimation method, should they note any non-normality prior to psychometric analysis. Previous research on the BSS-R has generally demonstrated benign and distributionally normal data characteristics and thus maximum-likelihood estimation without modification is invariably used for the CFA's undertaken. However, if non-normality is observed in future studies corrections such as the Satorra-Bentler (Satorra & Bentler, 2001) procedure for a scaled chi-square may be justified and legitimately undertaken and will likely improve model fit. The use of a correction such as Satorra-Bentler should be justified at the outset and on the premise of data characteristics rather than a post-hoc approach to fit improvement and as noted, in most validation studies of the BSS-R data is generally

distributionally normal. Using the total BSS-R score based on good fit of the bi-factor model has been suggested by Martin et al. (2018) and recommended by the ICHOM in the *Standard Set for Pregnancy and Childbirth* (Nijagal et al., 2018). The finding that the SE subscale, but not the BSS-R total score correlated with age, would also suggest using the total score should potential concerns arise regarding potential bias toward older mothers in terms of the SE sub-scale. All subscale scores and the total scale were found to have good internal consistency, endorsing this domain of validity and as Martin et al. (2018) indicate the use of the BSS-R as either a sub-scaled measure or the total score should be predicated by the purpose of application whether that be clinical, research or both.

It was observed that D-BSS-R sub-scale and sub-scale–total score correlations were for the most part significantly higher than those of the original UK-BSS-R (Hollins Martin & Martin, 2014), which may potentially be an insight into maternity practice in the Netherlands. A significant proportion of women in the current study enjoyed a home birth (22%), as well as the option for a birth centre that offers a known midwife and a maternity care assistant to support them during labor and birth. This profile of maternity service availability and uptake may contribute to the relatively increased (compared to UK) correlational relationship between D-BSS-R sub-scales. Indeed, the ‘Dutch Birth Centre Study’ demonstrated that client experiences of midwife-led care reduced the odds of suboptimal satisfaction on several measures (Hitzert et al., 2016). Furthermore, midwives job satisfaction in the Netherlands is generally high (Wiegers, Hermus, Verhoeven, Rijnders, & van der Pal-de Bruin, 2018), which may also work towards improving QC provided and reducing the stress experienced by childbearing women.

KGDV analysis showed an unsurprising finding that women who had an elective section experienced greater birth satisfaction, compared with those hurried for an emergency Caesarian Section. Similarly, women delivering a spontaneous vertex delivery experienced greater birth satisfaction than those who received assisted operative delivery (e.g., ventouse, forceps or Caesarean Section). Unsurprisingly and generally these findings match those of other international BSS-R studies (e.g., Fleming et al., 2016; Hollins Martin & Martin, 2014; Jefford et al., 2018; Nespoli et al., 2020). Data also shows very good convergent delivery with the DSM-5 derived PCL-5 questionnaire on all sub-scales, along with supporting a significant relationship between higher birth BSS-R sub-scale and total scores and reduced levels of trauma. This is both consistent and extends the findings of the Israeli BSS-R study (Skvirsky et al., 2019) and the large UK study (Harrison et al., 2020) to a validated measure of PTSD. The potential for birth satisfaction being a predictor of PP-PTSD would certainly be indicated as a future research avenue from the current study. Indeed, the BSS-R is a core instrument in the International Survey of Childbirth-Related Trauma (Ayers, 2021) study currently underway, thus future illumination of the relationship between birth experience and trauma is likely to produce unique insights into the dynamics of these associations in the near future and contextualised within salient cultural and service delivery characteristics. What is known is that up to 45% of women find childbirth traumatic, with 4% proceeding to develop *Post Traumatic Stress Disorder-Post Childbirth* (PTSD-PC) (J. Patterson, C.J. Hollins Martin, & T. Karatzias, 2019b). Also, perceptions of care provider's interpersonal behaviors are significantly associated with women developing psychological trauma (J. Patterson, C.J. Hollins Martin, & T. Karatzias, 2019a; Patterson et al., 2019b). For this reason, it is important to improve the intranatal

environment in ways that will enhance midwives' aptitude to interact well with women, which will in turn work towards improving QC and SE scores measured by the BSS-R.

Finally, it was noted that the D-BSS-R was a statistically significant predictor of the PCL-5 score in contrast to time since delivery which was not found to predict this score. This is an important observation since measures such as the BSS-R are retrospective instruments and the recall periods within a population can be from days to years. This finding would indicate that there is little impact of time since delivery on the relationship between birth experience assessed by the BSS-R and trauma score.

Conclusion

The current investigation has found the D-BSS-R to be a generally valid and reliable measure of birth experience consistent with the measurement characteristics of the original version of the tool. Conflicting findings were found across model fit indices in both the three-factor and bi-factor models which suggests further research with the measure is required to determine if future modification of the measure is required.

The study also corroborates previous suggestions of linkage between birth satisfaction and PP-PTSD using a robust and diagnostically valid measure of trauma. Further research is recommended with the measure in the Netherlands in order to confirm the measurement characteristics described in the current study, particularly in relation to clarification regarding model fit to data across model fit indices and investigate the utility of the measure in other groups of Dutch women.

Availability of the BSS-R

The BSS-R is free to use for clinical and research purposes but requires permission. Contact Professor Hollins Martin c.hollinsmartin@napier.ac.uk for permission to use and the dedicated dedicated BSS-R website www.bss-r.co.uk for more information on the measure.

Acknowledgements

We gratefully acknowledge the assistance of all those who took part in our study.

The authors are also extremely grateful for the comments, advice and suggestions of two expert reviewers on an earlier version of this manuscript. Their comments were both encouraging and thought-provoking in equal measure which made responding to the issues raised a pleasure.

References

- Alfaro Blazquez, R., Corchon, S., & Ferrer Ferrandiz, E. (2017). Validity of instruments for measuring the satisfaction of a woman and her partner with care received during labour and childbirth: Systematic review. *Midwifery*, *55*, 103-112. doi:10.1016/j.midw.2017.09.014
- Ayers, S. (2021). International Survey of Childbirth-Related Trauma (INTERSECT). Retrieved from <https://www.researchregistry.com/browse-the-registry#home/registrationdetails/5ffc7453702012001b80a58c/>
- Ayers, S., Bond, R., Bertullies, S., & Wijma, K. (2016). The aetiology of post-traumatic stress following childbirth: a meta-analysis and theoretical framework. *Psychological Medicine*, *46*(6), 1121-1134. doi:10.1017/S0033291715002706
- Ayers, S., & Pickering, A. D. (2001). Do women get posttraumatic stress disorder as a result of childbirth? A prospective study of incidence. *Birth*, *28*(2), 111-118. doi:10.1046/j.1523-536x.2001.00111.x
- 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), 504-511. doi:10.1080/02646838.2015.1024211
- Beck, C. T. (2004a). Birth trauma: in the eye of the beholder. *Nursing Research*, *53*(1), 28-35. doi:10.1097/00006199-200401000-00005
- Beck, C. T. (2004b). Post-traumatic stress disorder due to childbirth: the aftermath. *Nursing Research*, *53*(4), 216-224. doi:10.1097/00006199-200407000-00004

- Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, 107(2), 238-246. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/2320703>
- Boeschoten, M. A., Bakker, A., Jongedijk, R. A., & Olf, M. (2014). *PTSD checklist for the DSM-5 (PCL-5)—Nederlandstalige versie Uitgave*. Retrieved from Diemen, The Netherlands: Arq Psychotrauma Expert Groep:
- Boorman, R. J., Devilly, G. J., Gamble, J., Creed, D. K., & Fenwick, J. (2014). Childbirth and criteria for traumatic events. *Midwifery*, 30(2), 255-261. doi:10.1016/j.midw.2013.03.001
- Brown, T. (2015). *Confirmatory Factor Analysis for Applied Research* (2nd ed.). New York: Guilford Press.
- Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development *Psychological Assessment*, 7(3), 309-319.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297–334.
- Diedenhofen, B., & Musch, J. (2015). cocor: A comprehensive solution for the statistical comparison of correlations. *PloS One*, 10(4), e0121945. doi:10.1371/journal.pone.0121945
- 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, 9-15. doi:10.1016/j.midw.2016.07.008
- 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*, 1-11.

doi:10.1080/02646838.2018.1443322

Harrison, S. E., Ayers, S., Quigley, M. A., Stein, A., & Alderdice, F. (2020).

Prevalence and factors associated with postpartum posttraumatic stress in a population-based maternity survey in England. *Journal of Affective Disorders*.

doi:10.1016/j.jad.2020.11.102

Hitzert, M., Hermus, M. A., Scheerhagen, M., Boesveld, I. C., Wiegers, T. A., van

den Akker-van Marle, M. E., . . . de Graaf, J. P. (2016). Experiences of

women who planned birth in a birth centre compared to alternative planned

places of birth. Results of the Dutch Birth Centre Study. *Midwifery*, 40, 70-78.

doi:10.1016/j.midw.2016.06.004

Hodnett, E. D. (2002). Pain and women's satisfaction with the experience of

childbirth: a systematic review. *American Journal of Obstetrics and*

Gynecology, 186(5 Suppl Nature), S160-172. doi:10.1067/mob.2002.121141

Hollander, M. H., van Hastenberg, E., van Dillen, J., van Pampus, M. G., de Miranda,

E., & Stramrood, C. A. I. (2017). Preventing traumatic childbirth experiences:

2192 women's perceptions and views. *Archives of Womens Mental Health*,

20(4), 515-523. doi:10.1007/s00737-017-0729-6

Hollins Martin, C. J., & Martin, C. R. (2014). Development and psychometric

properties of the Birth Satisfaction Scale-Revised (BSS-R). *Midwifery*, 30(6),

610-619. doi:10.1016/j.midw.2013.10.006

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, 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), 42-58.

doi:10.1080/02646838.2017.1396302

Kline, P. (2000). *A Psychometrics Primer*. London: Free Association Books.

Kline, R. B. (2011). *Principles and Practice of Structural Equation Modeling* (3rd ed.).

London: Guilford Press.

Martin, C. R., Hollins Martin, C. J., Burduli, E., Barbosa-Leiker, C., Donovan-Batson, C., & Fleming, S. E. (2018). The Birth Satisfaction Scale - Revised (BSS-R):

should the subscale scores or the total score be used? *Journal of Reproductive and Infant Psychology*, 1-6.

doi:10.1080/02646838.2018.1490498

Mortazavi, F., Mehrabadi, M., Hollins Martin, C. J., & Martin, C. R. (2020).

Psychometric properties of the birth satisfaction scale-revised (BSS-R) in a sample of postpartum Iranian women. *Health Care for Women International*,

1-16. doi:10.1080/07399332.2020.1802464

Murray, A. (2013). The limitations of model fit in comparing the bi-factor versus

higher-order models of human cognitive ability structure. *Intelligence*, 41.

doi:10.1016/j.intell.2013.06.004

Nasiri, S., Kariman, N., & Ozgoli, G. (2020). Psychometric properties of the Iranian

version of Birth Satisfaction Scale-Revised. *Journal of Research in Medical Sciences*, 25(1), 90-90. doi:10.4103/jrms.JRMS_248_19

Nespoli, A., Colciago, E., Fumagalli, S., Locatelli, A., Hollins Martin, C. J., & Martin,

C. R. (2020). Validation and factor structure of the Italian version of the Birth

Satisfaction Scale-Revised (BSS-R). *Journal of Reproductive and Infant Psychology*, 1-16. doi:10.1080/02646838.2020.1836333

Patterson, J., Hollins Martin, C. J., & Karatzias, T. (2019a). Disempowered midwives and traumatised women: exploring the parallel processes of care provider interaction that contribute to women developing Post Traumatic Stress Disorder (PTSD) Post Childbirth. *Midwifery*, 76, 21-35. doi:10.1016/j.midw.2019.05.010

Patterson, J., Hollins Martin, C. J., & Karatzias, T. (2019b). PTSD post-childbirth: a systematic review of women's and midwives' subjective experiences of care provider interaction. *Journal of Reproductive and Infant Psychology*, 37, 56-83. doi:10.1080/02646838.2018.1504285

Romero-Gonzalez, B., Peralta-Ramirez, M. I., Caparros-Gonzalez, R. A., Cambil-Ledesma, A., Hollins Martin, C. J., & Martin, C. R. (2019). Spanish validation and factor structure of the Birth Satisfaction Scale-Revised (BSS-R). *Midwifery*, 70, 31-37. doi:10.1016/j.midw.2018.12.009

Satorra, A., & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika*, 66, 507-514.

Skodova, Z., Nepelova, Z., Grendar, M., & Baskova, M. (2019). Psychometric properties of the Slovak version of the Birth Satisfaction Scale (BSS) and Birth Satisfaction Scale-Revised (BSS-R). *Midwifery*, 79, 102550. doi:10.1016/j.midw.2019.102550

Skvirsky, V., Taubman-Ben-Ari, O., Hollins Martin, C. J., & Martin, C. R. (2019). Validation of the Hebrew Birth Satisfaction Scale - Revised (BSS-R) and its relationship to perceived traumatic labour. *Journal of Reproductive and Infant Psychology*, 1-7. doi:10.1080/02646838.2019.1600666

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. (2016). *Pregnancy and Childbirth Standard Set and Reference Guide*. Retrieved from <http://www.ichom.org/medical-conditions/pregnancy-and-childbirth/>

van Stenus, C. M., Boere-Boonekamp, M. M., Kerkhof, E. F., & Need, A. (2017). Client satisfaction and transfers across care levels of women with uncomplicated pregnancies at the onset of labor. *Midwifery*, 48, 11-17. doi:10.1016/j.midw.2017.02.007

Vardavaki, Z., Hollins Martin, C. J., & Martin, C. R. (2015). Construct and content validity of the Greek version of the Birth Satisfaction Scale (G-BSS). *Journal of Reproductive and Infant Psychology*, 33(5), 488-503. doi:10.1080/02646838.2015.1035235

Weathers, F. W., Litz, B. T., Keane, T. M., Palmieri, P. A., Marx, B. P., & Schnurr, P. P. (2013). *The PTSD Checklist for DSM-5 (PCL-5) – Standard [Measurement instrument]*. Retrieved from <https://www.ptsd.va.gov/>

Wiegers, T. A., Hermus, M. A., Verhoeven, C. J., Rijnders, M. E., & van der Pal-de Bruin, K. M. (2018). Job satisfaction of maternity care providers in the Netherlands: Does working in or with a birth centre influence job satisfaction? *European Journal of Midwifery*, 2, 11. doi:10.18332/ejm/94279

Williams, C., Patricia Taylor, E., & Schwannauer, M. (2016). A Web-Based Survey of Mother-Infant Bond, Attachment Experiences, and Metacognition in Posttraumatic Stress Following Childbirth. *Infant Mental Health Journal*, 37(3), 259-273. doi:10.1002/imhj.21564

Zafar, S., Tayyab, F., Liaqat, A., Sikander, S., Hollins Martin, C. J., & Martin, C. R. (2021). Translation and validation of the Birth Satisfaction Scale-Revised in Urdu for use in Pakistan. *International Journal of Childbirth*. doi:10.1891/ijcbirth-d-21-00001

Table 1. Mean, standard deviation and distributional characteristics of individual Dutch BSS-R items, sub-scale totals and the total Dutch BSS-R score. se = standard error of kurtosis.

Item	Item content	Domain*	Mean	SD	Min	Max	Skew	Kurtosis	se
BSS-R 1	I came through childbirth virtually unscathed	SE	2.33	1.43	0	4	-0.33	-1.28	0.09
BSS-R 2	I thought my labour was excessively long	SE	2.77	1.34	0	4	-0.75	-0.70	0.09
BSS-R 3	The delivery room staff encouraged me to make decisions about how I wanted my birth to progress	QC	2.70	1.30	0	4	-0.65	-0.76	0.08
BSS-R 4	I felt very anxious during my labour and birth	WA	2.34	1.29	0	4	-0.38	-0.94	0.08
BSS-R 5	I felt well supported by staff during my labour and birth	QC	3.28	1.04	0	4	-1.44	1.21	0.07
BSS-R 6	The staff communicated well with me during labour	QC	3.19	1.12	0	4	-1.33	0.82	0.07
BSS-R 7	I found giving birth a distressing experience	SE	2.30	1.37	0	4	-0.32	-1.18	0.09
BSS-R 8	I felt out of control during my birth experience	WA	2.35	1.49	0	4	-0.40	-1.29	0.10
BSS-R 9	I was not distressed at all during labour	SE	1.99	1.38	0	4	0.03	-1.30	0.09
BSS-R 10	The delivery room was clean and hygienic	QC	3.61	0.66	1	4	-1.60	1.89	0.04
Stress	Sub-scale total		9.39	4.14	0	16	-0.30	-0.90	0.27
Attributes	Sub-scale total		4.70	2.47	0	8	-0.40	-0.87	0.16
Quality	Sub-scale total		12.78	3.11	3	16	-1.03	0.35	0.20
Total	Total score		26.86	8.33	3	40	-0.48	-0.70	0.54

*Domain of the Dutch BSS-R. SE = Stress experienced during child-bearing, WA = Women's attributes, QC = Quality of Care.

Table 2. Confirmatory factor analysis of the Dutch BSS-R.

Model	χ^2 (df)	p	RMSEA	SRMR	CFI
1. Single factor	335.037 (35)	<0.001	0.189	0.102	0.738
2. Three-factor	112.786 (32)	<0.001	0.102	0.072	0.929
3. Bi-factor	81.81 (25)	<0.001	0.097	0.052	0.950

Table 3. Correlations of Dutch BSS-R sub-scales and total score and comparison with original UK BSS-R validation study (Hollins Martin and Martin, 2014).

Scale combination	Current study <i>r</i>	UK study <i>r</i>	Z	95% CI	<i>p</i>
Stress-Attributes	0.51	0.57	0.91	(-0.19 – 0.07)	0.36
Stress-Quality	0.49	0.26	2.90	(0.07 – 0.39)	0.004
Attributes-Quality	0.51	0.35	2.12	(0.01 – 0.31)	0.04
Total score-Stress	0.91	0.86	2.52	(0.01 – 0.09)	0.01
Total score-Attributes	0.87	0.80	2.52	(0.02 – 0.13)	0.01
Totals score-Quality	0.77	0.63	3.00	(0.05 – 0.24)	0.003

Table 4. Cronbach's alpha of Dutch BSS-R sub-scales and total score and comparison with original UK BSS-R validation study (Martin and Hollins Martin, 2014).

Degrees of freedom = 1.

Subscale	Current study	UK study	χ^2	p
Stress	0.74	0.71	0.42	0.52
Attributes	0.72	0.64	1.22	0.27
Quality	0.72	0.74	0.19	0.66
Total score	0.86	0.78	7.03	0.005

Table 5. Comparison of Dutch BSS-R total and sub-scale scores differentiated by birth delivery type. Standard deviations are in parentheses, degrees of freedom = 239, CI = confidence interval.

BSS-R Scale	Spontaneous vertex delivery (N=161)	Assisted/ Operative delivery (N=80)	95% CI	<i>t</i>	<i>p</i>	Hedges <i>g</i>	Hedges <i>g</i> 95% CI	Effect size
Stress	10.60 (3.75)	6.96 (3.85)	2.62 – 4.65	7.03	<0.001	0.96	0.68 – 1.24	Large
Attributes	5.38 (2.29)	3.33 (2.24)	1.44 – 2.67	6.60	<0.001	0.90	0.62 – 1.85	Large
Quality	13.53 (2.55)	11.26 (3.58)	1.48 – 3.06	5.65	<0.001	0.70	0.49 – 1.04	Medium
Total score	29.55 (7.25)	21.55 (7.85)	5.94 – 9.96	7.80	<0.001	1.06	0.78 – 1.35	Large

Table 6. Comparison of Dutch BSS-R total and sub-scale scores differentiated by Caesarean section type. Standard deviations are in parentheses, degrees of freedom = 43, CI = confidence interval.

BSS-R Scale	Elective (N=17)	Emergency (N=28)	95% CI	<i>t</i>	<i>p</i>	Hedges <i>g</i>	Hedges <i>g</i> 95% CI	Effect size
Stress	9.71 (4.25)	6.07 (3.28)	1.36 – 5.91	3.22	0.002	0.97	0.33 – 1.61	Large
Attributes	3.82 (2.40)	2.86 (2.30)	-0.49 – 2.42	1.34	0.19	0.41	-0.21 – 1.02	Small
Quality	13.47 (3.22)	10.04 (3.62)	1.28 – 5.59	3.22	0.002	0.97	0.33 – 1.61	Large
Total score	27.00 (8.27)	18.96 (7.29)	3.28 – 12.79	3.41	0.001	1.03	0.38 – 1.68	Large

Table 7. Correlation coefficients between the PCL-5 total score and Dutch BSS-R subscale score and total score.

Scale	PCL-5	Stress	Attributes	Quality	BSS-R total
PCL-5	1				
Stress	-0.45*	1			
Attributes	-0.48*	0.76*	1		
Quality	-0.35*	0.49*	0.45*	1	
BSS-R total	-0.50*	0.92*	0.85*	0.75*	1

* $p < 0.001$

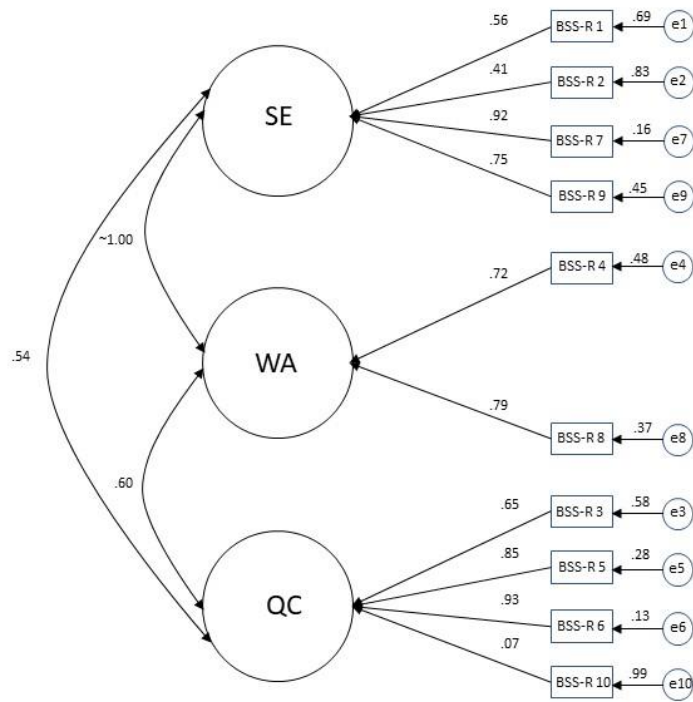


Figure 1. Standardised factor loadings of the tri-dimensional measurement model of the BSS-R. Correlations between factors and error term values are also indicated.