Validation and factor structure of the Italian version of the Birth Satisfaction Scale-Revised (BSS-R)

Antonella Nespoli[1]

Elisabetta Colciago[1*]

Simona Fumagalli[1]

Anna Locatelli[1,2]

Caroline J. Hollins Martin[3]

and

Colin R. Martin[4]

Affiliations:

[1]School of Medicine and Surgery, University of Milano - Bicocca, Via Cadore 48, 20900, Monza, Italy.

[2] Unit of Mother and Child, ASST Vimercate.

[3]School of Health and Social Care, Edinburgh Napier University (ENU), Sighthill Campus, Edinburgh EH11 4BN, Scotland, UK.

[4]Institute of Clinical and Applied Health Research (ICAHR), Faculty of Health Sciences, University of Hull, HU6 7RX, United Kingdom.

*Corresponding author: Elisabetta Colciago, School of Medicine and Surgery, University of Milano - Bicocca, Via Cadore 48, 20900, Monza, Italy. E-mail: e.colciago2@campus.unimib.it. ORCID: 0000-0002-2279-3641

Abstract

Objective: To validate the Italian-language version of the Birth Satisfaction Scale-Revised (BSS-R) and report key measurement properties of the tool. To evaluate the impact of antenatal class attendance on BSS-R assessed birth satisfaction.

Background: Maternal satisfaction is one of the standards of care defined by the World Health Organization (WHO) to improve the quality of services. The BSS-R is a multi-dimensional self-report measure of the experience of labour and birth.

Methods: Cross-sectional instrument evaluation design examining factor structure and key aspects of validity and reliability. Embedded between-subjects design to examine known-group discriminant validity and the impact of antenatal class attendance on BSS-R sub-scale and total scores as dependent variables. After giving birth, 297 women provided data for analysis.

Results: The Italian version of the BSS-R (I-BSS-R) was the key study measure. The established three-factor and bi-factor models of the BSS-R were found to offer an excellent fit to the data. Comparison of the tri-dimensional measurement model and the bi-factor model of the BSS-R found no significant differences between models. Women who attended antenatal classes had significantly lower stress experienced during childbearing sub-scale scores (I-BSS-R SE), compared to those who did not. Good convergent, divergent validity and known-groups discriminant validity was established for the I-BSS-R. Internal consistency observations were found to be sub-optimal in this population.

3

Conclusions: On all key psychometric indices, with the exception of internal consistency that requires further investigation, the I-BSS-R was found to be a valid translation of the original BSS-R. The impact of antenatal classes on birth satisfaction warrants further research.

Key words: Birth Satisfaction Scale-Revised (BSS-R), Birth satisfaction, Maternal satisfaction, Validation, Midwifery care**.**

Introduction

Giving birth is a complex psychological individual experience, with elements of universal physiological processes and life event significance (Larkin et al., 2009). Evidence suggest that the experience of labour and birth is complex and subjective (Larkin et al., 2009). A positive perceptions and satisfaction with the birth experience can be influenced by expectations' fulfilment, staff characteristics including quality of care and support, involvement in decision making, woman centred care and women's perception of control (Bayes et al., 2008; DeLuca & Lobel, 2014; Hildingsson, 2015; Hollander et al., 2017; Lewis et al., 2016). Furthermore, women's experience with birth could have long-term implications for woman and baby's health, both physically and emotionally (Karlström et al., 2015).

The WHO (World Health Organization, 2016) reported that satisfaction reflects the extent to which expectations of service standards have been met. There is consensus that satisfaction with care (Christiaens & Bracke, 2007; Larkin et al., 2009) is a complex psychological individual experience, with elements of universal physiological processes and life event significance, therefore influenced by a variety of factors. Satisfaction has been defined as a "positive feeling" or "affective response" to an event (Bramadat & Driedger, 1993). Understanding women's perception of care and satisfaction with services is important, as perceived quality is a key determinant of service utilisation (Srivastava et al., 2015). Healthcare systems could be more effective if they considered women's experiences, with the aim to provide quality care and meet families' needs and expectations (Chief Nursing Officers of England, Northen Ireland, Scotland and Wales, 2010; Rao et al., 2006).

Maternal satisfaction is one of the standards of care defined by the World Health Organization (WHO) to improve the quality of maternity services and to evaluate the

organization of Health Care Systems (World Health Organization, 2016); it should be considered as one of the most relevant indicators within both the midwifery and obstetric fields. Childbirth is one of the most common reason for accessing health facilities, therefore planners, managers and health care providers should assess women's satisfaction with care to evaluate services (Goodman et al., 2004; Hodnett, 2002).

A relatively under-explored though critically important element of care of relevance to the birth experience and birth satisfaction concerns the impact of antenatal classes. Contemporary practice and evidence advocates the use of antenatal classes to optimise birth preparedness to enhance the birth experience (Ricchi, A. et al., 2020), however this position is equivocal, with observations that the content of classes are not consistent with maternal expectations (Pålsson et al., 2019) and wide variation in the content of antenatal classes (Barimani et al., 2018).

Women's satisfaction with childbirth is an important measure of the quality of maternity care services, its assessment should be conducted using validated self-completion questionnaires due to their high reliability and low cost (Romero-Gonzalez et al., 2019). There are several instruments specifically developed to assess maternal satisfaction with care received during labour and birth (Alfaro Blazquez et al., 2017; Nilvér et al., 2017).

A systematic review (Nilvér et al., 2017) conducted with the aim to identify and present validated instruments measuring women's childbirth experience, collected 36 tools. Among these, two Scales have been used within the Italian context. "The childbirth perception questionnaire" by Bertucci et al. (Bertucci et al., 2012) does not present testing of psychometric properties and it should be further evaluated. The scale

6

"Women's delivery experience measures" by Mannarini et al. (Mannarini et al., 2013) aimed to evaluate birth experiences after both spontaneous and medically assisted pregnancy, focusing on indices considering the type of conception. This tool comprises a high number of items, 18, and does not focus only on the intrapartum care experience. Considering the Italian birth context and the model of midwifery care provided, the BSS-R was evaluated as the most appropriate instrument to be culturally validated, in order to evaluate maternal satisfaction with birth. The Italian maternal care is quite medicalised (Euro-Peristat Project, 2018), obstetricians are the primary providers of all antenatal care with the majority of women having a private doctor, who will not be present at their birth. Although the Italian National Healthcare System is free of change at the point of use, 44.7% of Italian women choose to pay for a private obstetrician (Lauria L, Lamberti A, Buoncristiano M, Bonciani M, Andreozzi S, 2012), even though they have the opportunity to go to a free public community or hospital. The lack of continuity of maternity care (Lauria L, Lamberti A, Buoncristiano M, Bonciani M, Andreozzi S, 2012), with the majority of women who get to know a midwife only at the time of labour and birth, concentrated our choice on a scale that could focus on the intrapartum care aspects. This could give the opportunity to assess the quality of the intrapartum midwifery care and to evaluate and implement quality improvement programs, in order to offer maternity services based on women's needs.

The Birth Satisfaction Scale-Revised (BSS-R; (Hollins Martin & Martin, 2014)) is a validated 10-item self-report measure that was developed in United Kingdom to evaluate women's satisfaction with birth. Comprising three sub-scales of : (i.) quality of care provided, (ii.) personal attributes of women and (iii.) stress experienced during childbirth. The BSS-R is a short, valid, reliable and theoretically-anchored measure to assess mothers' satisfaction with birth and has recently been recommended as the key

outcome measure for assessing birth experience globally (Nijagal et al., 2018). Widely translated and validated, the BSS-R has been shown to demonstrate generally excellent psychometric properties and conceptual alignment to the original UK version (Barbosa-Leiker et al., 2015; Burduli et al., 2017; Fleming et al., 2016; Jefford et al., 2018; Martin et al., 2017; Romero-Gonzalez et al., 2019; Vardavaki et al., 2015). The Italian version of the BSS-R has been recently developed following an extensive translation process to ensure congruence with the original version's conceptual and theoretical alignment, as well as emphasizing contextual anchoring in an Italian childbearing context (Nespoli et al., 2018). To date however, the Italian version of the BSS-R (I-BSS-R) has yet to be evaluated in a clinical population to determine and establish key psychometric properties, thus the current investigation sought to validate the I-BSS-R in relation to these measurement parameters. Consistent with the approach taken with other translation/validation studies of the BSS-R (e.g. (Romero-Gonzalez et al., 2019)) our objectives were to:

1.   Replicate the established tri-dimensional measurement model of the BSS-R in the I-BSS-R

2.   Assess the psychometric properties of the Italian version of the BSS-R

3.   Evaluate the impact of antenatal class attendance on I-BSS-R total and sub-scale scores.

Methods

Design

A cross-sectional instrument evaluation design utilising convenience sampling incorporating an embedded between-subjects design for known-groups discriminant validity testing. The Italian birth context has a classification system for levels of maternal care for Obstetric Units, comprising Level I Maternity Units providing care to women with low risk pregnancies or with minor complications and Level II Maternity Units dedicated to high risk conditions. Participants were recruited from an Obstetric Unit of a Level I Italian Maternity Hospital. Participating mothers signed a consent form, which informed them of the voluntary nature of their participation, about the aim of the study, the procedures and the confidentiality of data (anonymous codification). Participants were consented to take part in the study and completed the I-BSS-R prior to discharge from hospital and within 72 hours after birth.

Ethical approval

Ethical approval was obtained from the Hospitals' Ethical Review Board. Written informed consent was gained from all the participants.

Measures

The BSS-R comprises ten items scored using a five-point Likert type response format ranging from strongly agree to strongly disagree for each item. Higher scores on both the total score and the sub-scales equate to comparatively greater birth satisfaction. The three sub-scales: stress experienced during child-bearing (SE; 4 items); quality of care

(QC; 4 items); and women's attributes (WA; 2 items) assess distinct aspects of birth satisfaction and a total score can also be calculated to indicate overall birth satisfaction. Several translation/validation studies have indicated both the robust measurement characteristics of the BSS-R and conceptual alignment with the tri-dimensional measurement model of the tool (Barbosa-Leiker et al., 2015; Burduli et al., 2017; Jefford et al., 2018; Martin et al., 2017; Romero-Gonzalez et al., 2019).

A brief review of the translation process of the BSS-R into Italian

The exhaustive process taken to translating the BSS-R into Italian is described in detail by Nespoli et al. (Nespoli et al., 2018) The Italian version of the BSS-R was developed achieving the cross-cultural and conceptual equivalence of the English instrument (World Health Organization., 2016). Briefly, the process reported by Nespoli and colleagues (Nespoli et al., 2018) included a five-step forward translation, expert panel translation, back translation, pre-testing and cognitive interviewing to derive the final agreed version. A sample of 100 women were recruited to the study to determine understandability of the measure during the pre-testing and cognitive interviewing stages. This translation study required two BSS-R items to be modified; specifically item 1. '*I came through my childbirth unscathed*', was changed to '*I came through my childbirth without physical or psychological consequences*' and item 9. '*I was not distressed at all during labour*' was changed to '*I was not struggling at all during labour*'.

Statistical analysis

Confirmatory factor analysis

Evaluation of the tri-dimensional measurement model of the BSS-R (Objective 1) was conducted using Confirmatory Factor Analysis (CFA; (Brown, 2015)). CFA represents the statistical approach to evaluate a measurement model and has been applied to several BSS-R validation studies (Barbosa-Leiker et al., 2015; Göncü Serhatlıoğlu et al., 2018; Jefford et al., 2018; Vardavaki et al., 2015). The bi-factor model was also evaluated using CFA. The generally benign distributional characteristics of the BSS-R observed in previous studies (normal distribution, absence of skew and kurtosis) predicates a Maximum-likelihood approach to model estimation (R. B. Kline, 2011a). Consistent with conventional practice (Brown, 2015) multiple measures of model fit were applied, these being the comparative fit index (CFI: (Bentler, 1990), root mean squared error of approximation (RMSEA: (Byrne, 2010)), the square root mean residual (SRMR: (Hu & Bentler, 1999)), this also being consistent with contemporary BSS-R validation studies (e.g. (Jefford et al., 2018; Romero-Gonzalez et al., 2019).

It should be noted that in terms of model evaluation, $\chi2$ is influenced by both sample size and data variation, and contemporary practice is therefore to evaluate models using the fit indices above rather than $\chi2$ where sample size and trivial variations in data can lead to a significant $\chi2$ even with the context of well-fitting model (Byrne, 2010).

Divergent validity

Following the approach of Romero-Gonzalez (Romero-Gonzalez et al., 2019), divergent validity was assessed by correlation between the age of participants and I-BSS-R sub-scale and total scores with no statistically significant relationships between sub-scales/total score and age being anticipated.

Convergent validity

Taking a more sophisticated approach to convergent validity that has been seen in recent BSS-R validation studies to differentiate BSS-R sub-scales (e.g. (Romero-Gonzalez et al., 2019), convergent validity was assessed by examining the correlation between I-BSS-R sub-scales and the total I-BSS-R score and the number of days of gestation. It is predicted that the correlation between the I-BSS-R SE sub-scale and number of days gestation would be statistically significant and positive (thus greater satisfaction with duration of pregnancy). It is also predicted that there would be no statistically significant correlation between the I-BSS-R WA and I-BSS-R QC sub-scales and number of days gestation. No specific prediction is made regarding the relationship between the I-BSS-R total score and duration of pregnancy since this obviously represents a composite of all three sub-scales.

Known-groups discriminant validity

A standard approach taken with most BSS-R studies (e.g. (Jefford et al., 2018; Romero-Gonzalez et al., 2019; Vardavaki et al., 2015) to determine known-groups discriminant validity is to compare BSS-R sub-scale and total scores between those who have (i.) an unassisted vaginal delivery (UVD) and those who have an (ii.) assisted delivery (Intervention). An advantage of this approach to discriminant validity evaluation is the unambiguous discrete dichotomization of the between-groups variable that allows clarity in the evaluation of the discriminability of the measure and sub-scales along with the generation of effect sizes which can be used for comparison with observations from previous studies and the planning of future studies, for example in terms of sample size calculations. No instrument deliveries are offered within obstetric

practice at the research site therefore the intervention group comprises delivery by suction cap, elective Caesarean section or emergency Caesarean section. It is anticipated that those in the UVD group would have significantly higher I-BSS-R SE and WA sub-scale scores and I-BSS-R total scores compared to those in the Intervention group. BSS-R QC sub-scale scores vary between studies, with no difference on this sub-scale reported between groups in the original validation study (Hollins Martin & Martin, 2014) in contrast to large differences being observed in Greek, Spanish and United States BSS-R validation studies (Martin et al., 2017; Romero-Gonzalez et al., 2019; Vardavaki et al., 2015). Therefore, no specific prediction either in terms of statistical significance or directionality is made in relation to this I-BSS-R QC group comparison.

Antenatal class attendance

Antenatal class have the aim to guide parents toward childbirth and parenthood, providing evidence-based information. Antenatal classes start between 25 and 28 weeks of gestation, women decide spontaneously if attend them or not and they are suitable for first, second and third (plus) time parents. Classes are based either in hospital or in the community, in both cases they are run by a midwife and parents-to-be can choose which one they want to attend. Sometimes, during the hospital based antenatal classes, parents have also a meeting with an anesthetist, a pediatrician and an obstetrician. Classes are normally held once a week, for around two hours, for a total of 9 sessions during pregnancy and 1 following birth; usually these are pregnant women only classes with just a couple of partner sessions, when the midwife covers topics such as when to go to the hospital, what happens in labour and caring for the baby. Classes usually combine pregnancy-specific exercise (perlvic floor exercises, relaxation, positions during

labour), with information on different topics: physical and emotional changes during pregnancy, the development of the baby during the three trimesters, the stages of labour, coping strategies during childbirth, including different kinds of pain relief, feelings about birth, caring for a newborn baby, information on breastfeeding and parenthood. Antenatal class attendance dichotomized into (i.) Attended (A) and (ii.) Did Not Attend (DNA) was used as a categorical variable in a further evaluation of known-group discriminant validity. No specific predictions are made regarding differences in I-BSS-R sub-scales and the I-BSS-R total score between groups neither is directionality intimated.

Internal consistency

Internal consistency was evaluated using Cronbach's coefficient alpha (Cronbach, 1951) with I-BSS-R sub-scale and total score acceptability determined by threshold values of 0.70 or higher (P. Kline, 2000).

Results

Participants

Three-hundred Italian-speaking women took part in the study. Those consented to take part had given birth and completed the I-BSS-R within 72 hours postpartum. Three multivariate outliers were identified by reference to Mahalanobis distances exceeding the cut-off criteria of $\chi^2 > 29.59$ for a ten-item measure leaving a total sample size of N=297 for analysis. The mean age of participants was 33.12 (SD 4.96), range = 19-47 years. The mean duration of pregnancy was 39.36 (SD 1.16, range 37-41) weeks. One-hundred and sixty-seven (56%) participants had had their first baby. Comparison with the mean BSS-R sub-scale and total scores reported by Hollins Martin and Martin (2014) revealed all I-BSS-R sub-scales and the I-BSS-R total score to be significantly lower (Table 1.). Mean I-BSS-R sub-scale and total scores are summarised in Table 1. and reveal an absence of significant skew or kurtosis based on Kline's (R. B. Kline, 2011b) limits of 3 (skew) and 10 (kurtosis), a finding mirrored in the distributional characteristics of individual I-BSS-R items (Table 2.).

I-BSS-R SE, WA, and QC sub-scales were all highly correlated with the I-BSS-R total score, $r = 0.89$, $p <0.001$, $r = 0.73$, $p <0.001$ and $r = 0.47$, $p <0.001$, respectively. The SE sub-scale was observed to be significantly correlated with the WA ($r = 0.51$, $p <0.001$) and QC ($r = 0.19$, $p = 0.001$) sub-scales. No statistically significant correlation was observed between WA and QC sub-scales ($r = 0.05$, $p = 0.44$). Replicating the approach taken previously (Romero-Gonzalez et al., 2019), I-BSS-R sub-scale/total score correlation were compared with those reported by Hollins Martin and Martin (Hollins Martin & Martin, 2014) in the original BSS-R development study. Adopting the statistical approach of Diedenhofen and Musch (Diedenhofen & Musch, 2015)

statistically significant differences between studies were observed on two scale combinations, (i.) WA and QC sub-scales and (ii.) BSS-R total score and QC sub-scale. No other statistically significant differences were observed (Table 3).

Confirmatory factor analysis

The Kaiser-Meyer-Olkin (KMO) test of sampling adequacy and Bartlett's test of sphericity (BTS) indicated the suitability of the data for factor analysis, KMO = 0.71, BTS $\chi^2_{(df=45)}$ = 595.67, p < .001. Model fit of the three-factor measurement model of the BSS-R to data was observed to be excellent ($\chi^2_{(df=32)}$ = 52.52, p = 0.01, CFI = 0.96, RMSEA = 0.05, SRMR = 0.04). Evaluation of the bi-factor model of the BSS-R also revealed an excellent fit to the data ($\chi^2_{(df=25)}$ = 44.96, p = 0.01, CFI = 0.96, RMSEA = 0.05, SRMR = 0.05). Since the bi-factor model was found to offer an acceptable fit to the data, comparison with the three-factor correlated model using the chi-square differences test is deemed permissible (Reise, 2012; Yuan & Bentler, 2004). Comparison of the tri-dimensional measurement model with the bi-factor model revealed no statistical difference between models ($\Delta\chi^2$ = 7.56$_{(\Delta df=7)}$, p =.37).

Divergent validity

Correlations between the I-BSS-R total score, SE, WA and QC sub-scale scores and participant age were all non-significant, r = 0.004, p = 0.95, SE r = 0.02, p = 0.76, WA r = 0.02, p = 0.77, QE = -0.04, p = 0.46, respectively.

Convergent validity

Correlations between S-BSS-R total score and the SE sub-scale score and number of days gestation were both positive and statistically significant, $r = 0.17$, $p = 0.004$, $r = 0.21$, $p <0.001$, respectively. No statistically significant correlations were observed between WA and QC sub-scales and number of days gestation, $r = 0.02$, $p = 0.72$, $r = 0.08$, $p = 0.16$, respectively.

Known-groups discriminant validity

Highly statistically significant differences were observed on I-BSS-R total score and SE and QC sub-scales in favour (higher scores) of the UVD group compared to the intervention group. No difference was observed between groups on WA sub-scale scores (Table 4).

Participants who attended antenatal classes were observed to have significantly *lower* I-BSS-R SE sub-scale scores compared to those who did not attend such classes. No other statistically significant differences were observed (Table 5).

Internal consistency

Cronbach's alphas for the I-BSS-R total scale and sub-scales are summarised in Table 6. Significantly lower Cronbach's alpha's were observed on the I-BSS-R total scale and SE and QC sub-scales compared to those reported by Hollins Martin and Martin (Hollins Martin & Martin, 2014). No significant difference was observed between the I-BSS-R WA sub-scale and that reported by Hollins Martin and Martin (Hollins Martin & Martin, 2014).

Discussion

The current investigation sought to validate the recently developed Italian version of the BSS-R in an Italian clinical population. Across a range of psychometric criteria, the I-BSS-R appears to perform well and is consistent with the original BSS-R, however, on indices of internal consistency, findings are inconsistent with previous translation/validation studies of this increasingly used measure and therefore, overall, the current findings are equivocal. To understand the findings in both a theoretical and a clinically-applied manner, each of the key observations from the psychometric evaluation of the I-BSS-R will be examined on a point by point basis. Finally, the observation of significant differences in I-BSS-R SE sub-scale scores as a function of antenatal class attendance will also be explored.

Starting with the measurement model, the confirmatory factor analysis revealed an excellent fit to the tri-dimensional model of the BSS-R (Hollins Martin & Martin, 2014), which is a finding not only consistent with the original measure but also consistent with several other validation studies (Göncü Serhatlıoğlu et al., 2018; Jefford et al., 2018; Martin et al., 2017; Romero-Gonzalez et al., 2019; Vardavaki et al., 2015). The bi-factor model of the BSS-R was also found to offer an excellent fit to data and thus our findings are consistent with the bi-factor model of the BSS-R evaluated by Martin et al. (Martin et al., 2018) evidence of which supports the use of BSS-R sub-scale scores and the BSS-R total score, or both, depending on the clinical context of practice application of the measure and/or clinical research. Thus, in terms of routine clinical outcome monitoring the use of the BSS-R total score can be utilized with confidence, this being consistent with the recommended use of the measure and total item scoring approach recommended by The International Consortium for Health

Outcome Measurement (2016) Pregnancy and childbirth standard set guidelines. Drilling down further into the three sub-scales within the instrument, application to research in understanding the relationship of these domains of stress, attributes and quality of care offer the opportunity to explore these factors in terms of their relative sensitivity to discrete aspects of clinical context in terms of elucidating the specific areas of birth experience that impact on birth satisfaction and utilizing such observations to improve care and the woman's individual birth experience. Our study was also the first to empirically compare the established three-factor measurement model of the BSS-R to the bi-factor model and we found no statistically significant difference between models thus further supporting the assertion of Martin et al. (Martin et al., 2018) regarding equivalence of scoring approach.

Further validity evidence for the BSS-R was found in the divergent analysis observations which consistent with prediction confirmed the absence of significant relationships with participants' age, a finding consistent with the recent investigation by Romero-Gonzalez et al. (Romero-Gonzalez et al., 2019) and thus indicating no age-related adjustment is required to I-BSS-R scores.

In terms of novel insights from the application of the BSS-R in this population, the findings that the BSS-R SE sub-scale was highly correlated with duration of pregnancy (as was also the BSS-R total score), offers a useful insight into the little explored relationship between length of gestation and stress experienced during childbearing. Traditional focus has been around the issues of pre-term birth and primarily the health and well-being of the baby rather than the mother, for example Maslow, Caramanica, Li, Stellman, and Brackbill (2016). Our findings would suggest that further research is desirable into this relationship at term and specifically into conceptualizing the relationship as a continuum, thus illuminating other areas of clinical wellbeing of the

mother herself, which also are underexplored in the literature. For example, the timing of interventions at term. The findings do not intend to promote an induction of labour at term, which has the potential to do more harm than good, and its resource implications are staggering (Menticoglou & Hall, 2002; World Health Organization, 2011). Moreover, induction of labour, could also be perceived as another intrapartum intervention that leads to high level of stress with potentially a negative impact on maternal satisfaction. Instead, we suggest that women should be well informed and supported when they get close to the due date, in fact they would like and should be involved in decision-making to decide how to proceed once term is reached (Schwarz et al., 2016). These strategies could improve women's satisfaction with the entire birth process.

The finding that women experiencing an intervention delivery compared to an unassisted vaginal delivery have significantly lower I-BSS-R SE sub-scale and I-BSS-R total scores is entirely consistent with several BSS-R translation/validation studies that confirm the negative impact of intervention on birth satisfaction in relation to the BSS-R SE sub-scale and total BSS-R score (Hollins Martin & Martin, 2014; Jefford et al., 2018; Martin et al., 2017; Romero-Gonzalez et al., 2019; Vardavaki et al., 2015). Clearly this observation is not a critique of interventions *per se*, however it is an inditement of the negative impact of birth satisfaction and a pertinent reminder of the need to consider carefully the actual clinical need for an intervention (is the intervention *really* necessary?) and women's informed choice in the decision-making matrix regarding optimizing their birth experience.

The finding that women who attended antenatal classes had significantly lower BSS-R SE sub-scale scores compared to those who did not attend classes is striking in terms of both directionality and level of statistical significance. Antenatal classes are generally perceived as being of universal benefit for preparing women for birth and its potential consequences. Our study notes a negative impact from attending antenatal classes and effects upon maternal birth satisfaction. In accordance with Soriano-Vidal et al. (Soriano-Vidal et al., 2018), women who receive antenatal education are more likely to gain evidence-based information, which in turn increases their awareness and empowers them in relation to understanding birth processes. As such, receiving education could change or increase women's expectations of their birth experience, with mothers who have higher expectations experiencing lower fulfilment in relation to personal requests (Christiaens et al., 2008; Mei et al., 2016), which will effect overall birth satisfaction (Mei et al., 2016). Furthermore, we should consider the Italian context, where there is a high-risk culture surrounding childbirth (Euro-Peristat Project, 2018; Rota et al., 2017), which could add even more discrepancy between perceived best practice and actual midwifery care provided during labour and birth. Nonetheless, it is right that the midwifery services should strive to be evidence-based in their approach towards educating childbearing women, with knowledge imparted inevitably shaping women's birth experience.

The internal consistency observations represent one aspect of the current validation study that raises an area of concern and is inconsistent with the psychometric evaluation outlined thus far which generally indicates exemplary measurement characteristics. The Cronbach alpha findings of the I-BSS-R sub-scales and the total measure were all sub-optimal according to conventional criteria of 0.70 (P. Kline, 2000). This represents a surprising finding, not only in terms of the exhaustive translation process undertaken

by Nespoli et al. (Nespoli et al., 2018) to ensure equivalence of items to the original BSS-R, but also in terms of previous validation studies which generally find acceptable alpha for the BSS-R total scale (Burduli et al., 2017; Hollins Martin & Martin, 2014; Jefford et al., 2018; Romero-Gonzalez et al., 2019; Vardavaki et al., 2015). However, with the exception of Burduli et al. (Burduli et al., 2017), the BSS-R WA sub-scale is invariable observed to have an alpha of <0.70 largely interpreted as a function of the small number of items in this particular BSS-R sub-scale (N=2). There have also been occasions in validation studies of the BSS-R where the BSS-R QC sub-scale (Romero-Gonzalez et al., 2019) and the BSS-R SE sub-scale (Göncü Serhatlıoğlu et al., 2018) have been observed to have alpha <0.70. The observation of a sub-optimal alpha for a particular BSS-R sub-scale in these studies has generally been considered against a backdrop of generally excellent psychometric properties of the measure on other tests of validity and reliability and thus given a context which accepts this as a relatively minor shortcoming where it occurs. Further, given that the BSS-R can be used as a total score measure, the consistent observations prior to the current study that alpha for the total BSS-R scale is always >0.70 has minimized the potential issue of sub-optimal alpha in a specific BSS-R sub-scale. Thus, our observations in terms of internal validity represent a departure from previous observations and suggest further investigation of this aspect of reliability is required, and possibly the most appropriate method is a replication study in a further Italian-speaking population, in contrast to further revision of the scale.

A potential limitation of the study was that a comparison between the I-BSS-R and another measure of birth satisfaction was not made. Given that this could offer a valuable insight into both the I-BSS-R and indeed the measure to which it is compared in terms of evaluating degree of overlap and congruence, it is suggested that future

research with the I-BSS-R compare the degree of association of the measure with another birth satisfaction scale. The findings in relation to the relationship of antenatal classes and education to birth satisfaction has clearly been identified within the current study as an area to be 'unpacked' through further research. Given the variability in content and format in antenatal class provision both within and across countries, this would likely represent a complex research agenda to develop, however, the implications for care would indicate such effort is both worthwhile and indeed necessary.

Conclusion

Findings from the current investigation have found the I-BSS-R to be a valid and reliable version of the BSS-R, and for the most part consistent in terms of measurement characteristics and factor structure to the original English-language version. It was observed that women who attended antenatal classes had comparatively lower I-BSS-R SE sub-scale scores compared to antenatal class attendees a perhaps counterintuitive observation and one that clearly circumscribes an important agenda for further investigation in future research.

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. https://doi.org/10.1016/j.midw.2017.09.014

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. https://doi.org/10.1080/02646838.2015.1024211

Barimani, M., Forslund Frykedal, K., Rosander, M., & Berlin, A. (2018). Childbirth and parenting preparation in antenatal classes. *Midwifery*, *57*, 1–7. https://doi.org/10.1016/j.midw.2017.10.021

Bayes, S., Fenwick, J., & Hauck, Y. (2008). A Qualitative Analysis of Women's Short Accounts of Labour and Birth in a Western Australian Public Tertiary Hospital. *Journal of Midwifery & Women's Health*, *53*(1), 53–61. https://doi.org/10.1016/j.jmwh.2007.07.021

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, *107*(2), 238–246. https://doi.org/10.1037/0033-2909.107.2.238

Bertucci, V., Boffo, M., Mannarini, S., Serena, A., Saccardi, C., Cosmi, E., Andrisani, A., & Ambrosini, G. (2012). Assessing the perception of the childbirth experience in Italian women: A contribution to the adaptation of the childbirth perception questionnaire. *Midwifery*, *28*(2), 265–274. https://doi.org/10.1016/j.midw.2011.02.009

Bramadat, I. J., & Driedger, M. (1993). Satisfaction with Childbirth: Theories and Methods of Measurement. *Birth*, *20*(1), 22–29. https://doi.org/10.1111/j.1523-536X.1993.tb00175.x

Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (Second edition). The Guilford Press.

Burduli, E., Barbosa-Leiker, C., Fleming, S., Hollins Martin, C. J., & Martin, C. R. (2017). Cross-cultural invariance of the Birth Satisfaction Scale-Revised (BSS-R): Comparing UK and US samples. *Journal of Reproductive and Infant Psychology*, *35*(3), 248–260. https://doi.org/10.1080/02646838.2017.1310374

Byrne, B. M. (2010). *Structural equation modeling with AMOS: Basic concepts, applications, and programming* (2. ed). Routledge.

Chief Nursing Officers of England, Northen Ireland, Scotland and Wales. (2010). *Midwifery 2020: Delivering Expectation, 2010*.

Christiaens, W., & Bracke, P. (2007). Assessment of social psychological determinants of satisfaction with childbirth in a cross-national perspective. *BMC Pregnancy and Childbirth*, *7*(1), 26. https://doi.org/10.1186/1471-2393-7-26

Christiaens, W., Verhaeghe, M., & Bracke, P. (2008). Childbirth expectations and experiences in Belgian and Dutch models of maternity care. *Journal of Reproductive and Infant Psychology*, *26*(4), 309–322. https://doi.org/10.1080/02646830802350872

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*(3), 297–334. https://doi.org/10.1007/BF02310555

DeLuca, R. S., & Lobel, M. (2014). Diminished Control and Unmet Expectations: Testing a Model of Adjustment to Unplanned Cesarean Delivery: Adjustment to Cesarean Delivery. *Analyses of Social Issues and Public Policy*, *14*(1), 183–204. https://doi.org/10.1111/asap.12040

Diedenhofen, B., & Musch, J. (2015). cocor: A Comprehensive Solution for the Statistical Comparison of Correlations. *PLOS ONE*, *10*(4), e0121945. https://doi.org/10.1371/journal.pone.0121945

Euro-Peristat Project. (2018). *European Perinatal Health Report. Core indicators of the health and care of pregnant women and babies in Europe in 2015*. www.europeristat.com

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. https://doi.org/10.1016/j.midw.2016.07.008

Göncü Serhatlıoğlu, 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*, *36*(3), 235–245. https://doi.org/10.1080/02646838.2018.1443322

Goodman, P., Mackey, M. C., & Tavakoli, A. S. (2004). Factors related to childbirth satisfaction. *Journal of Advanced Nursing*, *46*(2), 212–219. https://doi.org/10.1111/j.1365-2648.2003.02981.x

Hildingsson, I. (2015). Women's birth expectations, are they fulfilled? Findings from a longitudinal Swedish cohort study. *Women and Birth*, *28*(2), e7–e13. https://doi.org/10.1016/j.wombi.2015.01.011

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.

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 Women's Mental Health*, *20*(4), 515–523. https://doi.org/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. https://doi.org/10.1016/j.midw.2013.10.006

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

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. https://doi.org/10.1080/02646838.2017.1396302

Karlström, A., Nystedt, A., & Hildingsson, I. (2015). The meaning of a very positive birth experience: Focus groups discussions with women. *BMC Pregnancy and Childbirth*, *15*(1). https://doi.org/10.1186/s12884-015-0683-0

Kline, P. (2000). *A psychometrics primer*. Free Association.

Kline, R. B. (2011a). *Principles and practice of structural equation modeling* (3. ed). Guilford Press.

Kline, R. B. (2011b). *Principles and practice of structural equation modeling* (3. ed). Guilford Press.

Larkin, P., Begley, C. M., & Devane, D. (2009). Women's experiences of labour and birth: An evolutionary concept analysis. *Midwifery*, *25*(2), e49–e59. https://doi.org/10.1016/j.midw.2007.07.010

Lauria L, Lamberti A, Buoncristiano M, Bonciani M, Andreozzi S. (2012). *Percorso nascita: Promozione e valutazione della qualità di modelli operativi. Le indagini del 2008-2009 e del 2010-2011. Roma: Istituto Superiore di Sanità; 2012. (Rapporti ISTISAN 12/39).* http://old.iss.it/binary/publ/cont/12_39_web.pdf

Lewis, L., Hauck, Y. L., Ronchi, F., Crichton, C., & Waller, L. (2016). Gaining insight into how women conceptualize satisfaction: Western Australian women's perception of their maternity care experiences. *BMC Pregnancy and Childbirth*, *16*(1). https://doi.org/10.1186/s12884-015-0759-x

Mannarini, S., Boffo, M., Bertucci, V., Andrisani, A., & Ambrosini, G. (2013). A Rasch-based dimension of delivery experience: Spontaneous vs. medically assisted conception. *Journal of Clinical Nursing*, *22*(17–18), 2404–2416. https://doi.org/10.1111/jocn.12264

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), e172–e178. https://doi.org/10.1016/j.wombi.2016.11.006

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*, *36*(5), 530–535. https://doi.org/10.1080/02646838.2018.1490498

Maslow, C. B., Caramanica, K., Li, J., Stellman, S. D., & Brackbill, R. M. (2016). Reproductive Outcomes Following Maternal Exposure to the Events of September 11, 2001, at the World Trade Center, in New York City. *American Journal of Public Health*, *106*(10), 1796–1803. https://doi.org/10.2105/AJPH.2016.303303

Mei, J. Y., Afshar, Y., Gregory, K. D., Kilpatrick, S. J., & Esakoff, T. F. (2016). Birth Plans: What Matters for Birth Experience Satisfaction. *Birth*, *43*(2), 144–150. https://doi.org/10.1111/birt.12226

Menticoglou, S. M., & Hall, P. F. (2002). Routine induction of labour at 41 weeks gestation: Nonsensus consensus. *BJOG: An International Journal of Obstetrics and Gynaecology*, *109*(5), 485–491.

Nespoli, A., Colciago, E., Pedroni, S., Perego, S., & Fumagalli, S. (2018). The Birth Satisfaction Scale-Revised (BSS-R): Process of translation and adaptation in an Italian context. *Annali Dell'Istituto Superiore Di Sanita*, *54*(4), 340–347. https://doi.org/10.4415/ANN_18_04_11

Nijagal, M. A., Wissig, S., Stowell, C., Olson, E., Amer-Wahlin, I., Bonsel, G., Brooks, A., Coleman, M., Devi Karalasingam, S., Duffy, J. M. N., Flanagan, T., Gebhardt, S., Greene, M. E., Groenendaal, F., R Jeganathan, J. R., Kowaliw, T., Lamain-de-Ruiter, M., Main, E., Owens, M., … Franx, A. (2018). Standardized outcome measures for pregnancy and childbirth, an ICHOM proposal. *BMC Health Services Research*, *18*(1). https://doi.org/10.1186/s12913-018-3732-3

Nilvér, H., Begley, C., & Berg, M. (2017). Measuring women's childbirth experiences: A systematic review for identification and analysis of validated instruments. *BMC Pregnancy and Childbirth*, *17*(1). https://doi.org/10.1186/s12884-017-1356-y

Pålsson, P., Kvist, L. J., Persson, E. K., Kristensson Hallström, I., & Ekelin, M. (2019). A survey of contemporary antenatal parental education in Sweden: What is offered to expectant parents and midwives' experiences. *Sexual & Reproductive Healthcare*, *20*, 13–19. https://doi.org/10.1016/j.srhc.2019.01.003

Rao, K. D., Peters, D. H., & Bandeen-Roche, K. (2006). Towards patient-centered health services in India—A scale to measure patient perceptions of quality. *International Journal for Quality in Health Care*, *18*(6), 414–421. https://doi.org/10.1093/intqhc/mzl049

Reise, S. P. (2012). The Rediscovery of Bifactor Measurement Models. *Multivariate Behavioral Research*, *47*(5), 667–696. https://doi.org/10.1080/00273171.2012.715555

Ricchi, A. et al. (2020). Study of childbirth education classes and evaluation of their effectiveness. *Clin Ter*, *170*(1), e78–e86. https://doi.org/doi: 10.7417/CT.2020.2193

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. https://doi.org/10.1016/j.midw.2018.12.009

Rota, A., Antolini, L., Colciago, E., Nespoli, A., Borrelli, S. E., & Fumagalli, S. (2017). Timing of hospital admission in labour: Latent versus active phase, mode of birth and intrapartum interventions. A correlational study. *Women and Birth*. https://doi.org/10.1016/j.wombi.2017.10.001

Schwarz, C., Gross, M. M., Heusser, P., & Berger, B. (2016). Women′s perceptions of induction of labour outcomes: Results of an online-survey in Germany. *Midwifery*, *35*, 3–10. https://doi.org/10.1016/j.midw.2016.02.002

Soriano-Vidal, F. J., Vila-Candel, R., Soriano-Martín, P. J., Tejedor-Tornero, A., & Castro-Sánchez, E. (2018). The effect of prenatal education classes on the birth

expectations of Spanish women. *Midwifery*, *60*, 41–47. https://doi.org/10.1016/j.midw.2018.02.002

Srivastava, A., Avan, B. I., Rajbangshi, P., & Bhattacharyya, S. (2015). Determinants of women's satisfaction with maternal health care: A review of literature from developing countries. *BMC Pregnancy and Childbirth*, *15*(1), 97. https://doi.org/10.1186/s12884-015-0525-0

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. https://doi.org/10.1080/02646838.2015.1035235

World Health Organization. (2011). *WHO recommendations for induction of labour.* World Health Organization.

World Health Organization. (2016). *Process of translation and adaptation of instruments.* Geneva.

World Health Organization. (2016). *Standards for improving quality of maternal and newborn care in health facilities*.

Yuan, K.-H., & Bentler, P. M. (2004). On Chi-Square Difference and z Tests in Mean and Covariance Structure Analysis when the Base Model is Misspecified. *Educational and Psychological Measurement*, *64*(5), 737–757. https://doi.org/10.1177/0013164404264853

Table 1. Mean, standard deviation and distributional characteristics of I-BSS-R sub-scales and total score and comparison with mean (Mu) UK BSS-R scores reported by Hollins Martin and Martin (2014) using one-sample *t*-test (df=296). se = standard error of kurtosis, CI = Confidence Interval.

Subscale	Mean	SD	Min	Max	Skew	Kurtosis	se	Mu	t	p
	95% CI									
Stress	8.28	3.12	0	16	-0.34	-0.05	0.18	9.70	7.86	<0.001
7.92 – 8.63										
Attributes	4.62	2.01	0	8	-0.34	-0.65	0.12	4.90	2.40	0.02
4.39 – 4.85										
Quality	14.07	1.69	8	16	-0.82	0.13	0.10	13.76	3.14	0.002
13.87 – 14.26										
Total Score	26.96	5.03	10	39	-0.29	0.12	0.29	28.36	4.78	<0.001
26.39 – 27.54										

Table 2. Mean, standard deviation and distributional characteristics of individual I-BSS-R items. se = standard error of kurtosis.

Item	Item content	Domain*	Mean	SD	Min	Ma
BSS-R 1	I came through my childbirth without physical or psychological consequences	SE	2.55	1.16	0	4
			2.49	1.24	0	4
BSS-R 2	I thought my labour was excessively long	SE	2.97	0.92	0	4
BSS-R 3	The delivery room staff encouraged me to make decisions about how I wanted my birth to progress	QC				
			2.27	1.24	0	4
BSS-R 4	I felt very anxious during my labour and birth	WA	3.73	0.50	2	4
BSS-R 5	I felt well supported by staff during my labour and birth	QC				
			3.70	0.55	1	4
BSS-R 6	The staff communicated well with me during labour	QC	1.79	1.17	0	4
BSS-R 7	I found giving birth a distressing experience	SE	2.35	1.14	0	4
BSS-R 8	I felt out of control during my birth experience	WA	1.43	1.03	0	4
BSS-R 9	I was not struggling at all during labour	SE	3.67	0.49	2	4
BSS-R 10	The delivery room was clean and hygienic	QC				

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

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

Scale combination	Italian r	UK r	Z	95% CI	p
Stress-Attributes	0.51	0.57	0.96	(-0.18 – 0.06)	0.34
Stress-Quality	0.19	0.26	0.83	(-0.23 – 0.10)	0.40
Attributes-Quality	0.05	0.35	3.56	(-0.46 – -0.14)	<0.001*
Total score-Stress	0.89	0.86	1.45	(-0.01 – 0.07)	0.15
Total score-Attributes	0.73	0.80	1.92	(-0.14 – 0.002)	0.06
Totals score-Quality	0.47	0.63	2.61	(-0.28 – -0.04)	0.009*

*p < 0.05.

Table 4. Comparison of I-BSS-R total and sub-scale scores as a function of birth delivery type. Standard deviations are in parentheses, degrees of freedom = 295, CI = confidence interval. Intervention group comprised Ceasarean section N=32, vacuum cup N=18.

BSS-R Scale Hedges g 95% CI	Unassisted Effect size* vaginal delivery (N=247)	Intervention delivery (N=50)	95% CI	t	p	Hedges g
Stress	8.55 (3.08)	6.90 (3.00)	0.72 – 2.59	3.48	<0.001	0.54
0.23 – 0.85	Medium					
Attributes	4.66 (2.03)	4.44 (1.92)	-0.40 – 0.83	0.69	0.49	0.11
-0.20 – 0.41	Negligible					
Quality	14.23 (1.56)	13.26 (2.03)	0.47 – 1.47	3.79	<0.001	0.59
0.28 – 0.90	Medium					
Total score	27.44 (4.95)	24.60 (4.79)	1.34 – 4.34	3.72	<0.001	0.58
0.27 – 0.88	Medium					

* Note: Effect size conventions, 0.2 = small, 0.5 = medium, 0.8 = large (Cohen, 1977).

Table 5. Comparison of I-BSS-R total and sub-scale scores as a function of attendance at antenatal classes. Standard deviations are in parentheses, degrees of freedom = 295, CI = confidence interval.

BSS-R Scale Hedges g 95% CI	Attended Effect size classes (N=152)	Did not attend classes (N=145)	95% CI	t	p	Hedges g
Stress	7.65 (3.19)	8.93 (2.92)	-1.98 - -0.58	3.60	<0.001	0.42
0.19 – 0.65	Small					
Attributes	4.70 (1.94)	4.53 (2.09)	-0.29 – 0.63	0.74	0.46	0.09
-0.14 – 0.31	Negligible					
Quality	14.08 (1.73)	14.06 (1.64)	-0.36 – 0.41	0.12	0.90	0.01
-0.21 – 0.24	Negligible					
Total score	26.43 (5.18)	27.52 (4.83)	-2.23 – 0.06	1.86	0.06	0.22
-0.01 – 0.44	Small					

Table 6. Cronbach's alpha of I-BSS-R sub-scales and total score and comparison with original UK BSS-R (Hollins Martin and Martin, 2014). Degrees of freedom = 1.

Subscale	Italian alpha	UK alpha	χ^2	p
Stress	0.60	0.71	3.90	0.05*
Attributes	0.60	0.64	0.23	0.63
Quality	0.57	0.74	9.41	0.002**
Total score	0.68	0.79	9.07	0.003**

*$p < 0.05$.

**$p < 0.005$.