

Hollins-Martin, C.J., Martin, C. (2014). Development and psychometric properties of the Birth Satisfaction Scale-Revised (BSS-R). *Midwifery*. 30: 610-619
<http://dx.doi.org/10.1016/j.midw.2013.10.006>

Caroline J. Hollins Martin¹ PhD MPhil BSc PGCE ADM RM RGN MBPsS

Colin R. Martin² PhD BSc RN YCAP CPsychol CSci AFBPsS

¹ Professor of Midwifery, School of Nursing, Midwifery and Social Care, University of Salford, Greater Manchester, England. Email: C.J.Hollins-Martin@salford.ac.uk, Telephone: 0161 2952 522

² Professor of Mental Health, School of Health, Nursing and Midwifery, University of the West of Scotland, Scotland, UK. Email: colin.martin@uws.ac.uk, Telephone: 01292 886336

Address for correspondence

Professor of Midwifery, School of Nursing, Midwifery and Social Care, University of Salford, Greater Manchester, England. Email: C.J.Hollins-Martin@salford.ac.uk, Telephone: 0161 2952 522

Development and psychometric properties of the Birth Satisfaction Scale-Revised (BSS-R)

Abstract

Objective: To assess factor structure, validity and reliability of the Birth Satisfaction Scale (BSS) and to develop a short-form version of the tool.

Design: A quantitative design focused on evaluating psychometric properties of the BSS using factor structure, internal consistency, divergent reliability and known groups validity.

Setting: Ayrshire Maternity Unit community midwife bases that serve the obstetric population of Ayrshire, Scotland (UK).

Participants: A convenience sample of healthy women (n=228) <10 days postpartum who had delivered a term infant. Data was collected from October 2010 to January 2011.

Measurement: The BSS contains 30 self report items, rated on a 5-point Likert scale that measure women's perceptions of: (1) *quality of care provision*, (2) *women's personal attributes*, and (3) *stress experienced during labour* (8, 8 & 14 items per factor).

Findings: Post data analysis the BSS was reconfigured into the 10 item BSS-Revised (BSS-R) comprised of 3 sub-scales that measure distinct but correlated domains of: (1) *quality of care provision*, (2) *women's personal attributes*, and (3) *stress experienced during labour*. These domains now consist of relatively few items (4, 2 & 4 items per factor), but offer a good fit to the data.

Key conclusions: The BSS-R would appear to be a robust, valid and reliable multi-dimensional psychometric instrument for measuring postnatal women's birth satisfaction. Further research to confirm the veracity of the instruments measurement properties highlighted in the current study is desirable. The BSS-R is available for use at a national/international level from the first author.

Key words: assessment, audit, Birth Satisfaction Scale (BSS), Birth Satisfaction Scale-Revised (BSS-R), childbearing, intranatal, labour, measurement, midwifery, validation.

Development and psychometric properties of the Birth Satisfaction Scale-Revised (BSS-R)

Introduction

The National Service Framework for Children, Young People and Maternity Services (DH, 2004) vitally focuses upon the needs of childbearing women. Stated within is a government pledge to enhance choice and control for childbearing women, with emphasis firmly placing her at the centre of care provision. These points are re-emphasised in “*Our health, our care, our say*” (DH, 2006) and “*Maternity matters: choice, access and the continuity of care in a safe service*” (DH, 2007). Every woman’s perceptions of birth are important, which within this study is conceptualised as “birth satisfaction”. In terms of quantitative research, a woman’s satisfaction with intranatal care can only be considered high quality when gratification over what she received is measured as high (Mahon, 1996). Meaningful measurement of such perceptions will only be achieved using a rigorous, valid and reliable psychometric instrument. To achieve this goal, Hollins Martin and Fleming (2011) developed a 30 item psychometric scale to assess women’s perceptions of birth called the Birth Satisfaction Scale (BSS). The BSS was developed to facilitate researchers, maternity care staff and consumers to construct a meaningful picture of what constitutes women’s like or dislike of the childbearing experience (Hollins Martin and Fleming, 2011).

Through undertaking a literature review and transcribing research-based expressions of birth satisfaction/dissatisfaction into statements, the items of the BSS were developed. Three overarching themes were identified (see *Table 1*).

TABLE 1 HERE

The aim of the present study was to determine the factor structure, validity and reliability of the BSS and to consider the realisability of a short-form version of the tool. A further aim was to reflect on the approaches taken by both Hollins Martin and Fleming (2011) in the development of the BSS and the work of Hollins Martin et al. (2012) in application of concurrent analysis to support the veracity of the underlying themes supporting construction of this measure.

Concurrent analysis

Hollins Martin et al. (2012) set about examining the parsimony of the BSS using a qualitative technique called concurrent analysis, which involves domain analysis to offer alternative explanations and assist better understanding of the importance of each item. Using the primary free text data collected from 19 qualitative papers the domain analysis confirmed 3 explanatory items within the BSS: *being in control*, *things going as planned*, and *being supported*. The conclusion drawn was that the BSS accounts for all the analysed data, suggesting it is a robust measure of satisfaction in childbirth. To view processes involved see Hollins Martin et al. (2012).

In summary, the properties of the BSS evaluated in the current study include factor structure, internal consistency, divergent reliability, and known groups validity, for purposes of answering the following research questions:

- (1) Is the BSS a uni-dimensional or multidimensional measure?
- (2) Are the thematically embedded and postulated sub-scales robust and reliable?
- (3) Does the BSS and any inherent sub-scales demonstrate acceptable internal consistency and divergent reliability?
- (4) Is the known groups validity of the BSS satisfactory?

(5) Is the 30-item BSS the most appropriate and psychometrically valid formulation of the tool?

(6) Can the BSS be revised to a more concise and robust measure?

Methods

Design

Utilising a prospective cross-sectional design, this study focused on evaluating key psychometric properties of the BSS, including a sequential optimisation of the instrument. This study follows a sequential process of instrument evaluation using classical and contemporary psychometric approaches (Byrne, 2010; Kline, 2000) applied to a single cohort that is differentiated by clinical attributes and which allowed evaluation of instrument measurement properties. Ethics approval was provided by the National Health Service (NHS) National Research Ethics Service (NRES), with data collected between October 2010 and January 2011.

Participants

Participants were a convenience sample of postnatal women in the first 10 postnatal days of receiving care from maternity service providers in the West of Scotland (UK). Those with a medical diagnosis, poor obstetric history, prematurity (< 37 weeks), postmaturity (> 42 weeks), < 16 years or > 50 years old were excluded from participating in the study. Medical diagnosis included women with pre-existing heart disease, pre-eclampsia, kidney problems, autoimmune disorders, cancer or severe infection. Poor obstetric history embraced having previous unfavourable fetal outcome in terms of intrauterine fetal death, intrauterine growth retardation, stillbirth, early neonatal death and/or congenital anomalies, all of which have the potential to compromise psychological well-being during labour. The community midwife in

attendance collected the data within the participant's home. Inclusion criteria embraced having had an uncomplicated pregnancy and delivering at term.

Data collection

Participants were recruited by the community midwife who provided routine postnatal care. Training in advance of data collection was specified at the community midwives office by the first author. An information sheet, consent form and the questionnaire were supplied in a pack to the woman by the community midwife at the first postnatal visit. A question and answer session was offered, with the completed questionnaire collected at the final appointment. The researcher's phone number and an address were made available to answer participants' subsequent queries and receive delayed completions. The first author in conjunction with the community manager coordinated processes.

Participants responded to the 30 BSS items on a 5-point Likert scale based on level of agreement or disagreement with each of the statements placed. Half of the items were reverse scored with a possible range of scores between 30-150. A score of 30 on the BSS represents least "birth satisfaction" and 150 most. The items comprising the BSS are shown in *Table 2*. An example question is shown in *Figure 1*.

TABLE 2 HERE

FIGURE 1 HERE

Statistical analysis

The objectives of the present study in terms of evaluation and potential optimisation of the BSS scale required the use of Confirmatory Factor Analysis (CFA) (Kline, 1993, 2000) and Structural Equation Modelling techniques (SEM) (Byrne, 2010; Kline, 1998, 2005) used in combination. It should be noted that factor analysis (both CFA and SEM) are considered a part of the multiple general linear hypothesis set of procedures and consequently share many of the fundamental assumptions of associated statistical techniques. These parametric assumptions include multivariate normality and normal distribution, absence of outliers and interval level of measurement. The robustness of parametric tests against violations of the fundamental parametric assumptions (Martin and Thompson, 2000) have resulted in the contemporary use of ordinal or ordered categorical data, which represents the common reality of questionnaire data, with these statistical techniques (Colman et al., 1997; Friedrich et al., 2011; Kind and Barmby, 2011; Shulruf et al., 2008). However, data exhibiting significant deviation from the normal distribution assumption can lead to an erroneous outcome of a statistical analysis based on assumed parametric acceptable data distributional characteristic and consequently, an incorrect and potentially misleading interpretation of statistical findings (Flora and Curran, 2004; Lubke and Muthen, 2004; Martin and Thompson, 2000; Muthen and Kaplan, 1992). Therefore, each of the BSS items distributional characteristics were examined in detail and evaluated to determine deviation from assumed normality which could have a deleterious impact on CFA and SEM. Skew and kurtosis characteristics of each item were examined and those exhibiting any significant deviation from normality were rejected from the BSS item pool prior to further statistical analysis based on normality assumptions. The criteria for item rejection based on univariate

skew and kurtosis characteristics and was based on absolute skew values equal to, or greater than 3 and absolute kurtosis values of equal to, or greater than 10, based on the non-normality cut-off recommendations of Kline (2005). Statistical analysis for all quantitative studies in this study was conducted using PASW version 18 (SPSS, 2009a,b), Analysis of Moment Structures (AMOS) version 18 (Arbuckle, 1995-2009) and Mplus version 3 (Muthen and Muthen, 1998-2004).

Structural Equation Modelling (SEM)

Evaluation of a measurement model of a psychometric measure can be conducted using SEM. The approach taken was entirely consistent with the special case modelling associated with SEM and CFA. However, it should be emphasised that though the approach taken is justified by the robustness of CFA approaches to model evaluation, by definition the use of this method to determine optimal scale and sub-scale structure necessarily defines this approach as fundamentally exploratory rather than confirmatory (Byrne, 2010). Legitimately, following the determination of the optimal item content for each sub-scale, a CFA can be conducted on an a priori specified model comprising three related thematically determined sub-scales of *quality of care provision, women's personal attributes and stress experienced during labour*, consistent with the themes identified in the original study by Hollins Martin and Fleming (2011). This would thus represent the a priori model and since this model is circumscribed within the original paper by Hollins Martin and Fleming, legitimately represents an a priori specified model, based on the evidence furnished by the instrument developer's thematic analysis.

The uni-dimensional model of the BSS, as intimated as a possible representative model of measure by Hollins Martin and Fleming (2011) can also be evaluated using the same approach, and indeed, SEM allows the competing models

to directly compare to determine the best fit to the data. Since a key goal within SEM is pursuit of model fit and parsimony, the model providing the best statistical fit to the data represents the most appropriate representation of the measurement model, the caveat being that model fit is statistically adequate.

Consistent with the assumption of multivariate normality, a maximum-likelihoods (ML) approach to model estimation was adopted (Byrne, 2010; Kline, 1993, 2000). Multiple goodness of fit tests (Bentler and Bonett, 1980) were used to evaluate the models, these being the comparative fit index (CFI) (Bentler, 1990) and the Root Mean Squared Error of Approximation (RMSEA). A CFI greater than 0.90 indicates an acceptable fit to the data (Bentler and Bonett, 1980; Bollen, 1989; Hu and Bentler, 1995; Kline, 1998; Marsh et al., 1988) while a CFI equal to or greater than 0.95 indicates a good fit to the data (Hu and Bentler, 1999). A RMSEA with values of less than 0.08 indicate an acceptable fit to the data (Browne and Cudeck, 1993) while values of less than 0.05 indicate a good fit to the data (Schumaker and Lomax, 2010). A statistically significant χ^2 indicates a significant proportion of variance within the data is unexplained by the model (Bentler and Bonett, 1980). However trivial and inconsequential variations in the data can promote a significant χ^2 likelihood test statistic (Hu and Bentler, 1995), hence model evaluation is almost universally determined by model fits statistics such as CFI and RMSEA (Byrne, 2010; Hooper et al., 2008). One approach to address the limitations of the χ^2 likelihood test statistic is to use the ratio of χ^2 to degrees of freedom (df) a model fit index (χ^2/df) with ratios of 2:1 indicative of a good model fit (Carmines and Mclver, 1981). The root mean square residual (RMR) represents a useful model fit index which assesses the discrepancy between the predicted model correlation matrix and sample data.

Standardised RMR (SRMR) values of 0.05 or less indicate good model fit (Byrne, 2010).

Model respecification

Unsatisfactory model fit to the three-factor and single factor models specified by SEM requires redress by further exploratory analysis based on the first principles of factor analysis. This involves examination of item-factor loadings and potential reduction of the number of items in the measure due to poor item-factor loadings or item-factor cross-loadings, which would allow the process of respecifying the measurement model and evaluating the veracity of the measurement model of the thematically-circumscribed BSS. This process was performed on the full 30-item BSS.

Determination of an indicative item-factor loading was set at 0.30 initially, consistent with other investigators who have used factor analysis (Hazlett-Stevens et al., 2004; Jomeen and Martin, 2004; Karimova and Martin, 2003; Martin and Thompson, 1999; Martin and Thompson, 2000). To optimise the process and help achieve a parsimonious solution, a cut-off of 0.40 (Upton and Upton, 2006) was used to determine the plausibility of the a priori and thematically-specified three-factor model. Item-factor loadings were also evaluated using their individual critical ratios, which should exceed 1.96 to be statistically significant. Finally, problematic or poorly performing items within the measurement model were identified by examining item-item standardised residual covariances, and improvement to model fit was assessed via modification indices.

Divergent validity

Divergent validity was determined by correlating BSS scale scores with the mother's age. It was predicted that there would be no significant relationship between BSS scores and the mother's age.

Known groups validity

Known groups validity was evaluated by testing for differences in BSS scores in relation to birth type and comparing normal to non-normal delivery (forceps, ventouse, prearranged section, emergency section). It was predicted that BSS total scores would be significantly higher for normal compared to non-normal delivery. In relation to theoretically circumscribed sub-scales embedded within the instrument, it is anticipated that while there would be no significant difference in *quality of care provision* sub-scale scores and *women's personal attributes* sub-scales scores as a function of delivery type. However, it was predicted that scores on the *stress experienced during labour* sub-scale would be significantly higher (since higher scores relate to greater satisfaction within this domain) in the normal delivery type group.

Internal consistency

An internal consistency analysis of the BSS was conducted to ensure that the measures satisfied the criteria for clinical and research purposes using the Cronbach coefficient alpha statistical procedure. A Cronbach's alpha reliability statistic of 0.70 is considered as the minimum acceptable criterion of instrument internal reliability (Kline, 1993; 2000).

Findings

Descriptive results

One-thousand two-hundred study information packs were given out of which 228 women consented to take part in the study and completed the BSS. This represents a response rate of 19%. One-hundred and ten (48%) of the participants were primigravidas. The average duration of pregnancy was 40.08 (SD 1.26) weeks. The average duration of labour was 8.60 (SD 6.81) hours with a reported range of 0-53 hours duration. The total BSS score was 115.84 (SD 14.05), and the thematically determined sub-scale mean scores of the *quality of care provision* (8 items), *women's personal attributes* (8 items), and *stress experienced during labour* (14 items) sub-scales, were 32.20 (4.09), 31.90 (4.16) and 51.73 (8.04) respectively. Details of birthing unit, type of delivery and pain relief are shown in *Table 3*.

TABLE 3 HERE

Multivariate normality

The distribution of the 30 BSS items revealed no significant evidence of skew or kurtosis, with the exception of BSS item 6. '*I gave birth to a normal healthy baby*' (skew = 3.32, kurtosis = 14.62). The mean, standard deviation, skew and kurtosis of each BSS item are summarised in TABLE 4. The mean score of the 30 item BSS revealed distributional normality as indicated by the Kolmogorov-Smirnov statistic with Lilliefors significance correction (K-S = 0.05, df = 228, $p=0.20$).

TABLE 4 HERE

Evaluation of predicted models

The structure of the thematically derived three-factor model was found to be poor ($\chi^2_{(df = 402)} = 1599.25$, $p < 0.001$, $\chi^2/df = 3.98$, CFI = 0.47, RMSEA = 0.11, RMR = 0.14 and SRMR = 0.11). The alternative single-factor model was also revealed to have a poor fit to the data ($\chi^2_{(df = 405)} = 1664.01$, $p < 0.001$, $\chi^2/df = 4.11$, CFI = 0.44, RMSEA = 0.12, RMR = 0.13 and SRMR = 0.11).

Respecification of measurement model of the BSS

Using the respecification protocol outlined previously, 20 BSS items were revealed to perform inadequately within the context of the a priori three-factor thematically derived structure. A three-factor model comprising correlated factors of *quality of care provision* (3 items), *women's personal attributes* (3 items), and *stress experienced during labour* (4 items) resulting in a 10-item scale was found to offer an adequate fit to the data ($\chi^2_{(df = 32)} = 70.47$, $\chi^2/df = 2.20$, $p < 0.001$, CFI = 0.93, RMSEA = 0.08, RMR = 0.05 and SRMR = 0.07). A single-factor model comprising these 10 items revealed a comparatively poorer fit to the data ($\chi^2_{(df = 35)} = 192.23$, $p < 0.001$, $\chi^2/df = 5.49$, CFI = 0.71, RMSEA = 0.15, RMR = 0.07 and SRMR = 0.11). The differences test revealed the superiority of the three-factor model compared to the single factor model to be statistically significant ($\chi^2_{diff(df = 3)} = 121.76$, $p < 0.001$). However, further review of this model through evaluation of modification indices suggested that BSS item 2 '*The delivery room staff encouraged me to make decisions about how I wanted my birth to progress*' would improve the fit of the model when specified as loading on the *quality of care provision* domain in contrast to loading on the *women's personal attributes* domain. The resulting model offered an excellent fit to the data ($\chi^2_{(df = 32)} = 42.67$, $p = 0.10$, $\chi^2/df = 1.33$, CFI = 0.98, RMSEA = 0.04, RMR = 0.04 and SRMR = 0.05) with a

non-significant χ^2 value suggesting little unexplained variance inherent within the model. The individual items comprising the 10-item model of the BSS and associated factor domains are shown in *Table 5*. The three-factor 10-item measurement model of the BSS-Revised (BSS-R) is shown in *Figure 2*. The mean score distribution of the 10 item BSS-R revealed some evidence of deviation from distributional normality as indicated by the Kolmogorov-Smirnov statistic with Lilliefors significance correction (K-S = 0.06, df = 228, $p = 0.02$). Boxplot analysis revealed no evidence of outlying or extreme score, while examination of the Q-Q plot suggested the data followed closely a normal distribution. Scrutiny of the frequency histogram of BSS-R scores also followed a relatively normal distribution, though there was evidence of a narrow secondary peak of scores in the 32-33 score range. In view of the high correlation between the domains of *stress experienced during labour* and *women's personal attributes*, the possibility that a two-factor model may offer a more parsimonious explanation of the data was explored by evaluating the model of fit of a combined *stress experienced during labour/women's personal attributes* factor correlated with the *quality of care provision* factor. The model offered a relatively good fit to the data (χ^2 (df = 34) = 53.85, $p = 0.02$, $\chi^2/df = 1.58$, CFI = 0.97, RMSEA = 0.05, RMR = 0.05 and SRMR = 0.06), though a significant proportion of variance remained unexplained by the model. However, comparison between best-fit two and three-factor model revealed the three-factor model to offer a significantly better fit to the data, (χ^2 diff_(df = 2) = 11.18, $p < 0.01$).

FIGURE 2 HERE

TABLE 5 HERE

Evaluation of a second-order model

A final adaptation of the best-fit CFA model of the 10-item BSS-R was evaluated based on the original work of Hollins Martin and Fleming (2011) which specifies that the BSS would be scored as a total score, thus a summary statistic of underlying birth satisfaction supporting the use of the instrument as a uni-dimensional tool. To examine this possibility within the context of the short-form instrument and extrapolating back to the total score recommendations of the instrument developers, a second-order CFA model was evaluated with a higher order domain of 'experience of childbirth', thus investigating whether the BSS-R is more appropriately realised as a hierarchical factorial structure with this domain explaining the first order factors. The resulting hierarchical model offered an excellent fit to the data ($\chi^2_{(df = 33)} = 44.41, p = 0.09, \chi^2/df = 1.35, CFI = 0.98, RMSEA = 0.04, RMR = 0.04$ and $SRMR = 0.05$) with a non-significant χ^2 value suggesting little unexplained variance inherent within the model. The model is shown in *Figure 3*.

FIGURE 3 ABOUT HERE

Rescaling of the BSS-Revised (BSS-R)

The revised version of the BSS (BSS-R) was rescaled to produce a zero score point across the total scale and the three sub-scales, thus the revised 10-item instrument would be scored along a 0-4 Likert scale. Using this approach the BSS-R total score was 28.36 (SD 5.78, range 13-40), and the thematically determined sub-scale mean scores of the *quality of care provision* (BSS-QC), *women's personal attributes* (BSS-WA), and *stress experienced during labour* (BSS-SL) sub-scales were 13.76 (SD 2.13, range 6-16), 4.90 (SD 1.98, range 0-8) and 9.70 (SD 3.29, range 0-16)

respectively. All three sub-scales were observed to be moderately correlated, with the relationship between these sub-scales measures shown in *Table 6*.

TABLE 6 HERE

Divergent validity

No significant correlation was observed between the BSS-R total score and mother's age ($r = 0.07$, $p=0.28$). Neither was there evidence of any significant relationships with the sub-scale scores: between BSS-SL scores, BSS-WA scores, BSS-QC and mother's age ($r = 0.07$, $p=0.27$; $r = 0.01$, $p=0.96$; $r = 0.08$, $p=0.21$); respectively.

Known groups validity

The mean BSS-R total score and BSS-SL, BSS-WA and BSS-QC sub-scale scores as a function of delivery type are shown in *Table 7*. Five women who had a pool birth were excluded from the analysis, since this does not typically represent a non-normal birth. Non-normal categorisation (N=81) included forceps and ventouse delivery, pre-organised and emergency caesarean section and breech birth. A significant difference between groups differentiated by the delivery type was observed on BSS-R total score ($t_{(221)} = 3.44$, $p = 0.001$), and BSS-SL sub-scale score, ($t_{(221)} = 4.81$, $p < 0.001$) in the direction predicted. No statistically significant difference was observed in BSS-QC sub-scale scores ($t_{(221)} = 0.24$, $p = 0.81$), as a function of delivery type. Finally, a statistically significant difference was observed in BSS-WA scores, with women in the normal delivery group have higher scores on this sub-scale ($t_{(221)} = 2.39$, $p = 0.02$).

TABLE 7 HERE

Internal consistency

Calculated Cronbach's alpha of the BSS-R total scale and BSS-SL, BSS-WA , BSS-QC sub-scales were 0.79, 0.71, 0.64 and 0.74 respectively.

Discussion

Findings from this enquiry suggest that the BSS-R is a robust, valid and reliable multi-dimensional psychometric instrument for measuring postnatal women's birth satisfaction. It is comprised of three sub-scales that measure distinct but correlated domains of: (1) *quality of care provision*, (2) *women's personal attributes*, and (3) *stress experienced during labour*. Observably, these domains consist of relatively few items (4, 2 & 4 items per factor). Nevertheless, they have been observed to offer a good fit to the data and provide a sound psychometric basis for assessing women's satisfaction with their birth experience. The original additional 20 items on the BSS did not demonstrate satisfactory levels of statistical validity and so were removed, which trimmed the questionnaire from a 30 to a 10 item questionnaire, with a possible range of scores lying between 0 and 40 (0 representing least satisfaction and 40 most) using a 0-4 scaling per item.

Findings from the statistical analysis at a surface level of appraisal, appear to contradict the validity findings of the Hollins Martin et al. (2012) concurrent analysis study, which concluded that the 30 item BSS accounted for all the analysed data and as such promoted it as a robust measure of birth satisfaction. However, it is important to note that the statistical approach used in the current study represented a quantitative focus on construct validity and reliability, whereas Hollins Martin et al. (2012) used a primarily qualitative approach to evaluate face/content validity. Thus, at a deeper level of sympathetic comparison, both the approaches and aims were not directly comparable. Indeed, the statistical results indicate a position that concurrent

analysis was primarily a validity assessment of the literature review that underpinned the initial 30 BSS item development (Hollins Martin and Fleming, 2011). However, what may be irreconcilable in terms of reflection on the empirical findings from the current study, is that concurrent analysis may indeed be a reinforcing tautology of the evidence that underpins item construction. Further, the notion that concurrent analysis is bound within tautological constraints is made as observation without agenda and without a positivity or negativity valence.

The updated valid and reliable 10 item scale produced from this study has been renamed the BSS-Revised (BSS-R) and the items are as follows:

(Q1) I came through childbirth virtually unscathed

Having an instrumental intervention or caesarean section is inextricably linked to receiving an obstetric injury (Beck, 2009; McKenzie-McHarg, 2004; Geissbuehler and Eberhard 2002; Sorenson and Tschetter, 2010), with depressed scores on the Edinburgh Postnatal Depression Scale from some of those in reception of having received an operative delivery and higher weighted scores for obstetric procedures (Brown and Lumley, 1994). Induced labour is also associated with more pain and ultimately less birth satisfaction (Faure, 1991; Niven and Gijsbers, 1984).

(Q2) I thought my labour was excessively long

Lengthy labour is a birth satisfaction indicator, with its increase in reports of pain and dissatisfaction with the experience (Niven and Gijsbers, 1984; Scott-Palmer and Skevington, 1981). Having a long labour is potentially compounded by parity, related obstetric factors, anxiety and reduced personal control (Faure, 1991; Niven and Gijsbers, 1984; Scott-Palmer and Skevington, 1981).

(Q3) The delivery room staff encouraged me to make decisions about how I wanted my birth to progress

Providing choice profoundly effects women's experiences of labour and birth (Gibbens and Thomson, 2001; Hall and Holloway, 1998; Handfield and Bell, 1995; Too, 1996), which necessitates that maternity care professionals empower women with knowledge to become constructive decision-makers (Melender, 2002) in relation to how they would like their labour to be managed (Berg et al., 1996; Brown and Lumley, 1994; Hodnett, 1996; Halldorsdottir and Karlsdottir, 1996; Walker et al., 1995).

(Q4) I felt anxious during my labour and birth

Anxious women have a predisposition towards having negative birth expectations (Heaman et al., 1992; Sjogren, 1997), which renders midwives accountable for instilling confidence through delivering effective preparatory education. Preparation for childbirth palpably influences birth satisfaction (Dannenbring et al., 1997), with engagers more confident and able to cope (Sinclair, 1999). It also enhances self-efficacy (Handfield and Bell, 1995), which has been shown to reduce pain experience (Larsen et al., 2001; Stockman and Altmaier, 2001).

(Q5) I felt well supported by staff during my labour and birth

Hodnett (2002) meta-analysed 14 trials (n=5,020 women) that measured the effects of continuous support from caregivers during labour on childbearing women's labour experience. Continuous support was associated with reduced requests for pain relief. Numbers receiving operative vaginal delivery and caesarean section were also reduced. In general, women who felt supported viewed their birth experience more favourably (6 trials) (Hodnett, 2002).

(Q6) The staff communicated well with me during labour

Quality of care provision is multifaceted. Aspects of relationships with staff are important (Grabowska, 2009; Proctor, 1998; Sorenson and Tschetter, 2010), with quality relationships including being offered information from which to make choices in plain English, with staff honest and consistent about what is provided (Proctor, 1998). Staff also require to be flexible, informal, interested and friendly (Chen, 2001; Proctor, 1998; Waldenstrom, 1998) at the same time as being professional, skilled and knowledgeable (Proctor, 1998). In the presence of negative birth perceptions and perceived low-quality relationships with providers, women's long-term memories of negative experiences can be preserved (Stadlmayr et al., 2006). When women feel treated as objects, such disaffirming is significantly correlated with negative birth perceptions (Sorenson and Tschetter, 2010). Negative outcomes are related to disaffirmation expressed through verbal and nonverbal provider interactions, with healing effects reported from supportive interactions (Bedell et al., 2004; Kirkpatrick et al., 2005).

(Q7) I found giving birth a distressing experience

Distress experienced during labour affects birth experience (Alehagen et al., 2000), with this inextricably linked to receiving an obstetric injury (Beck, 2009; McKenzie-McHarg, 2004; Geissbuehler and Eberhard 2002; Sorenson and Tschetter, 2010), such as caesarean section (Bryanton et al., 2008; Wax et al., 2004).

(Q8) I felt out of control during my birth experience

Feeling in control has been securely linked with women's experiences of birth satisfaction (Goodman et al., 2004; Green et al., 2003; Melender, 2002).

(Q9) *I was not distressed at all during labour*

Amount and type of pain experienced during labour is a birth satisfaction indicator (Quine et al., 1993), with primigravidas experiencing greater pain than multiparous women (Faure, 1991; Melzack et al., 1984; Niven and Gijbers, 1984).

(Q10) *The delivery room was clean and hygienic*

The environment is associated with making birth a more satisfying experience (Proctor, 1998; Waldenstrom, 1998). For most women, cleanliness is an essential feature, with uncontaminated delivery rooms considered safe places where infection risks are minimised (Proctor, 1998).

Both statistical analysis and the evidence supports that the 3-factor 10-item measurement model of the BSS-R shown in *Figure 2* can now be considered a psychometrically robust instrument for measuring women's satisfaction with their birth experience. To view the restructured BSS-R see *Table 8*.

TABLE 8 HERE

The statistical approach taken to the study also enjoyed an economy over more traditional two-stage exploratory factor analysis-CFA approaches (Kline, 2000) which require two separate samples, by monopolising on the underlying multidimensional model suggested in Hollins Martin and Fleming's (2012) original paper, thus requiring a single-stage, single-group analysis. It is statistically legitimate and appropriate to use the statistical approaches more readily identified within the CFA literature while acknowledging that the paucity of *empirical* data previously available on the BSS defines the approach to statistical illumination and

determination of the underlying factor structure of the instrument as exploratory. Indeed, it is the approach and rationale underpinning the statistical evaluation of the BSS within the context of the current study as exploratory rather than the actual method of statistical analysis *per se*. This does not imply that one approach is any way superior over the other, simply that the approach taken in the current investigation was justified on the basis of the multidimensional model suggested by Hollins Martin and Fleming (2012) and the theoretical underpinnings that supported the instruments multidimensionality from their scale development research.

An important caveat to the original study on the development of the BSS by Hollins Martin and Fleming (2011) concerns these researchers claim that the BSS is quintessentially a uni-dimensional measure, a claim implicit to the scoring approach devised (a single total score). The thrust of Hollins Martin and Fleming's (2011) original paper emphasised the tool as envisaged as a uni-dimensional construct and the three domains that emerged within their thematic review may appear, on reflection, to sit to a degree as a somewhat awkward anachronism arising from the paper, a tension within the original work. Arguably, the findings from the current study support the multi-domain perspective, specifically in terms of the poor fit of the single factor model of the 30-item BSS when applied to the current sample. A potential ramification of the need to reduce the scale to 10 items resulting in an improved fit offers and indication that the uni-dimensional conceptualisation of the full 30-item BSS represents a case of model miss-specification. Postulating this hypothesis further, retrospectively revisiting the original BSS development paper offers and indication that the balance within that paper may have been out of balance, with the emphasise on the uni-dimensionality of the scale being overstated and the tri-dimensional thematically-derived model being under-developed in relation

to exposition and theoretical and conceptual reach. It is clearly beyond the scope of the current study to determine this issue with any certainty beyond that of informed and reflective conjecture. However, raising this issue now as an element of relevant and dynamic concern to the future application and development of the BSS in either 30-item or 10-item guises, suggests a potential direction for further research and development of the scale.

Additionally, a note of caution is required regarding the use and interpretation of the BSS-R as a total score. It should be emphasised that model fit for a uni-dimensional model was poor and the final tri-dimensional model of the BSS-R offered an excellent fit to the data, thus supporting the multi-dimensionality of the tool. However, evidence was furnished for a higher second-order latent factor of *experience of childbearing*, underpinning quality, stress and attributes factors, offering an excellent fit to the data. A problem manifests however in that the veracity of the evidence to support the investigation of the higher order model is largely an artefact of a tension in the original literature reporting the BSS and the application of a single total score, with the suggestion that the tool may be multidimensional from being thematically-derived. Therefore the suggestion of a hierarchical model of the BSS-R, supported by good statistical model fit, offers the opportunity and justification supported from this analysis to investigate further the use of the BSS-R as a total score while also furnishing a context to explore further the underlying structure of the tool through additional empirical research. Consequently, our findings of a hierarchical model of the BSS-R facilitate the future evaluation of this instrument to determine what additional benefits and clinical utility, beyond those of the three subscales, may be found when the total score is used.

The modest response rate deserves consideration in terms of possible impact on findings and indeed, the generalisability of the findings from the current study. The response rate in terms of information packs distributed compared to consent and study participation was around 1 in 5. It is possible that those individuals who chose to participate in the study may differ in some characteristic way from those that did not participate which may impact on the response to the questionnaire. This represents a limitation of the current study and this issue may meaningfully be addressed by the availability of more BSS data as the instrument becomes used more extensively in a broader range of groups, where the impact of clinical and cultural factors may also be evaluated and reflected upon.

Finally, a further area of future research concerns a limitation in the current study in relation to convergent validity. The vacuum in the literature which predicated the development of the original BSS by Hollins Martin and Fleming (2011) also revealed a difficulty in identifying the most appropriate measure with which to determine the convergent validity of the scale. This represents an important omission in a comprehensive validation of the BSS and BSS-R and a legitimate and pressing focus of additional research enquiry with the current two versions of the tool and future evolution and translations of the scale.

Conclusion

Data has shown that the validated BSS-R is a robust tool for midwives, obstetricians and maternity care managers to measure postnatal women's birth satisfaction. In terms of impact, the BSS can be requested for use by researchers to collect data both nationally and internationally, with results potentially correlated with other measures (e.g., pain and/or depression scales). It has already been requested for use in several countries (America, Greece, Ireland, South Korea, Scotland) and where necessary was translated into the relevant language. If you would like to use the BSS-R then please contact the first author.

Acknowledgements

The researchers would like to thank Geraldine Butcher, the community managers and midwives who helped this study achieve what was intended. Without their efforts data collection would not have been possible. We are also extremely grateful to two referees for their very helpful and facilitative comments on two earlier versions of this manuscript. Their expert feedback was incredibly detailed and specific which helped make attending to the revisions both a pleasure and usefully reflective. We would also like to thank the 228 postnatal women who gave up their time to complete the BSS.

References

- Alehagen, S., Wijma, B. 2000. Can women's cognitive appraisal be registered throughout childbirth? *Gynecologic and Obstetric Investigation* 49, 31-5.
- Arbuckle, J.L. 1995-2009. *AMOS 18 User's Guide*. AMOS Development Corporation, Chicago, IL.
- Beck, C.T. 2009. Birth trauma and its sequelae. *Journal of Trauma and Dissociation* 10(2), 189-203.
- Bedell, S.E., Graboys, T.B., Bedell, E., Lown, B. 2004. Words that harm, words that heal. *Archives of Internal Medicine* 164(13), 1365-1368.
- Bentler, P.M. 1990. Comparative fit indexes in structural models. *Psychological Bulletin* 107, 238-246.
- Bentler, P.M., Bonett, D.G. 1980. Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin* 88, 588-606.
- Berg, M., Lundgren, I., Hermansson, E. 1996. Women's experience of the encounter with the midwife during childbirth. *Midwifery* 12, 11-15.
- Bollen, K.A. 1989. A new incremental fit index for general structural equation models. *Sociological Methods and Research* 17, 303-316.
- Browne, M.W., Cudeck, R. 1993. Alternative ways of assessing model fit. In K.A. Bollen and J.S. Long (Eds.), *Testing Structural Equation Models* (pp. 136-162). Sage, Newbury Park CA.
- Brown, S., Lumley, J. 1994. Satisfaction with care in labour and birth: a survey of 790 Australian women. *Birth* 21(1), 4-13.
- Bryanton, J., Gagnon, A., Johnston, C., Hatem, M. 2008. Predictors of women's perceptions of the childbirth experience. *Journal of Obstetric, Gynecologic and Neonatal Nursing* 37(1), 24-34.
- Byrne, B.M. 2010. *Structural equation modeling with AMOS: Basic concepts, applications, and programming* (2nd ed.). Routledge/Taylor and Francis Group, New York.
- Carmines, E.G., McIver, J.P. 1981. Analyzing models with unobserved variables: analysis of covariance structures. In: Bohrnstedt GW, Borgatta EF, editors. *Social measurement*. Thousand Oaks: Sage Publications, pp. 65-115.
- Chen, C., Wang, S., Chang, M. 2001. Women's perceptions of helpful and unhelpful nursing behaviours during labour: a study of Taiwan. *Birth* 28(3), 180-185.
- Colman, A. M., Norris, C. E., Preston, C. C. 1997. Comparing rating scales of different lengths: Equivalence of scores from 5-point and 7-point scales.

Psychological Reports 80(2), 355-362.

Dannenbring, D., Stevens, M.J., House, A.E. 1997. Predictors of childbirth pain and maternal satisfaction. *Journal of Behavioural Medicine* 20(2), 127-42.

Department of Health (DH) 2004. *The National Service Framework for Children, Young People and Maternity Services*. DH, London.

Department of Health (DH) 2006. *Our health, our care, our say*. DH, London.

Department of Health (DH) 2007. *Maternity Matters: choice, access and the continuity of care in a safe service*. DH, London.

Faure, E.A.M. 1991. The pain of parturition. *Semin Perinatol* 15, 342-7.

Flora, D.B., Curran, P.J. 2004. An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. *Psychological Methods* 9(4), 466-491.

Friedrich, F., Alexandrowicz, R., Benda, N., Cerny, G., Wancata, J. 2011. The criterion validity of different versions of the General Health Questionnaire among non-psychiatric inpatients. *Social Psychiatry and Psychiatric Epidemiology* 46(7), 635-641.

Geissbuehler, V., Eberhard, J. 2002. Fear of childbirth during pregnancy: a study of more than 8000 pregnant women. *Journal of Psychosomatic Obstetrics and Gynecology* 23, 229-35.

Gibbens, J., Thomson, A. 2001. Women's expectations and experiences of childbirth. *Midwifery* 17, 302-13.

Goodman, P., Mackay, M.C., Tavakoli, A.S. 2004. Factors related to childbirth satisfaction. *Journal of Advanced Nursing* 46, 212-9.

Grabowska, C. 2009. Unhappiness after childbirth. In C. Squire (Ed.), *The Social Context of Birth*. Radcliffe Publishing Ltd, Abingdon.

Green, J.M., Baston, H.A., Easton, S.C, McCormick, F. 2003. Greater expectations: the inter-relationship between women's expectations and experience of decision-making, continuity, choice and control in labor, and psychological outcomes, summary report. *Mother and Infant Research Unit (University of Leeds), Leeds*.

Halldorsdottir, S., Karlsdottir, S.I. 1996. Journeying through labour and delivery: perceptions of women who have given birth. *Midwifery* 12, 48-61.

Hall, S.M., Holloway, I.M. 1998. Staying in control: women's experiences of labour in water. *Midwifery* 14, 30-6.

Handfield, B., Bell, R. 1995. Do childbirth preparation classes influence decision making about labour and postpartum issues? *Birth* 22(3), 153-60.

Hazlett-Stevens, H., Ullman, J.B., Craske, M.G. 2004. Factor structure of the Penn State Worry Questionnaire: examination of a method factor. *Assessment* 11(4), 361-370.

Heaman, M., Beaton, J., Gupton, A. 1992. A comparison of childbirth expectations in high risk and low risk pregnant women. *Clinical Nursing Research* 1(3), 252-65.

Hodnett, E.D. 1996. Nursing support of the labouring woman. *Journal of Obstetric, Gynecological and Neonatal Nursing* 25(3), 257-64.

Hodnett, E.D. 2002. Review: continuous caregiver support during labour has beneficial maternal and infant outcomes. *Evidence-based Nursing* 5, 105.

Hollins Martin, C.J., Fleming, V. 2011. The Birth Satisfaction Scale (BSS). *International Journal of Health Care Quality Assurance* 24(2), 124-135.

Hollins Martin, C.J., Snowden, A., Martin, C.R. 2012. Concurrent analysis: Validation of the domains within the Birth Satisfaction Scale. *Journal of Reproductive and Infant Psychology* 30(3), 247-260.

Hooper, D., Coughlan, J., Mullen, M.R. 2008. Structural equation modelling: Guidelines for determining model fit. *Electronic Journal of Business Research Methods* 6(1), 53-60.

Hu, L.T., Bentler, P.M. 1995. Evaluating model fit. In R.H. Hoyle (Ed.), *Structural equation modelling: concepts, issues and applications*. Sage. Thousand Oaks, CA.

Jomeen, J., Martin, C.R. 2004. Is the hospital anxiety and depression scale (HADS) a reliable screening tool in early pregnancy? *Psychology and Health* 19(6), 787-800.

Karimova, G., Martin, C.R. 2003. A psychometric evaluation of the Hospital Anxiety and Depression Scale during pregnancy. *Psychology, Health and Medicine* 8, 89-103.

Kind, P., Barmby, P. 2011. Defending attitude scales. In M. Russell and M. Kavanaugh (eds.), *Assessing students in the margins: Challenges, strategies, and techniques*. (pp. 117-135). IAP Information Age Publishing, Charlotte NC US.

Kirkpatrick, J.N., Nash, K., Duffy, T.P. 2005. Well rounded. *Archives of Internal Medicine* 165(6), 613-616.

Kline, P. 1993. *The handbook of psychological testing*. Routledge, London.

Kline, R.B. 1998. *Principles and practice of structural equation modeling*. Guilford, New York.

Kline, P. 2000. *A psychometrics primer*. Free Association Books, London.

Kline, R.B. 2005. *Principles and Practice of Structural Equation Modeling*, 2nd ed, Guilford Press, New York

Larsen, K., O'Hara, M., Brewer, K., Wenzel, A. 2001. A prospective study of self-efficacy expectancies and labor pain, *Journal of Reproductive and Infant Psychology*. 19(3),203-14.

Lubke, G.H., Muthén, B.O. 2004. Applying Multigroup Confirmatory Factor Models for Continuous Outcomes to Likert Scale Data Complicates Meaningful Group Comparisons. *Structural Equation Modeling* 11(4), 514-534.

Mahon, P.Y. 1996. An analysis of the concept 'patient satisfaction' as it relates to contemporary nursing care. *Journal of Advanced Nursing* 24,1241-8.

Marsh, H.W., Balla, J.R., McDonald, R.P. 1988. Goodness-of-fit indices in confirmatory factor analysis: the effect of sample size. *Psychological Bulletin* 103, 391-410.

Martin, C.R., Thompson, D.R. 1999. Utility of the Hospital Anxiety and Depression Scale in patients with end-stage renal disease on continuous ambulatory peritoneal dialysis. . *Psychology, Health and Medicine* 4(4), 369-376.

Martin, C.R., Thompson, D.R. 2000. *Design and Analysis of Clinical Nursing Research Studies*. Routledge, London.

McKenzie-McHarg, K. 2004. Traumatic birth: understanding predictors, triggers, and counseling process is essential to treatment. *Birth* 31(3), 219-221.

Melender, H.L. 2002. Fears and coping strategies associated with pregnancy and childbirth in Finland. *Journal Midwifery Womens Health* 47, 256-63.

Melzack, R., Taenzer, P., Feldman, P., Kinch, R. 1984. Labor is still painful after prepared childbirth. *Canadian Medical Journal* 125, 357-63.

Muthen, B., Kaplan, D. 1992. A comparison of some methodologies for the factor analysis of non-normal Likert variables: A note on the size of the model. *British Journal of Mathematical and Statistical Psychology* 45(1), 19-30.

Muthen, L.K., Muthen, B.O. 1998-2004. *Mplus User's Guide*, 3rd ed. Muthen and Muthen, Los Angeles.

Niven, C., Gijsbers, K. 1984. A study of labor pain using the McGill pain questionnaire. *Social Science Medicine* 19, 1347-51.

Proctor, S. 1998. What determines quality in maternity care? Comparing the perceptions of childbearing women and midwives *Birth*, 25(2), 85-93.

Quine, L., Rutter, D.R., Gowan, S. 1993. Women's satisfaction with the quality of the birth experience: a prospective study of social and psychological predictors. *Journal of Reproductive and Infant Psychology* 11, 107-13.

Schumacker, R.E., Lomax, R.G. 2010. *A beginner's guide to structural equation*

modelling, 3rd ed. New York, Routledge/Taylor and Francis Group.

Scott-Palmer, J., and Skevington, S.M. 1981. Pain during childbirth and locus of control. *Journal of Psychosomatic Research* 25, 151-5.

Shulruf, B., Hattie, J., Dixon, R. 2008. Factors affecting responses to Likert type questionnaires: introduction of the ImpExp, a new comprehensive model. *Social Psychology of Education* 11(1), 59-78.

Sinclair, M. 1999. The childbirth self-efficacy inventory: a replication study. *Journal of Advanced Nursing* 30(6), 1416-23.

Sjogren, B. 1997. Reasons for anxiety about childbirth in 100 pregnant women. *Journal of Psychosomatic Obstetrics and Gynecology* 18, 266-72.

Sorenson, D.S., Tschetter, L. 2010. Prevalence of negative birth perception, disaffirmation, perinatal trauma symptoms, and depression among postpartum women. *Perspectives in Psychiatric Care* 46(1), 14-25.

SPSS. 2009a. PASW Statistics 18 Core System User's Guide. SPSS Inc., Chicago.

SPSS. 2009b. PASW Advanced Statistics 18. Chicago, IL: SPSS, Inc.

Stadlmayr, W., Amsler, F., Lemola, S., Stein, S., Alt, M., Bürgin, D., et al. 2006. Memory of childbirth in the second year: the long term effect of a negative birth experience and its modulation by the perceived intranatal relationship with caregivers. *Journal of Psychosomatic Obstetrics and Gynecology* 27(4), 211-224.

Stockman, A., Altmaier, E. 2001. Relation of self-efficacy to reported pain and pain medication usage during labour. *Journal of Clinical Psychology in Medical Settings* 8, 161-6.

Too, S.K. 1996. Do birth plans empower women? A study of their views. *Nursing Standard* 10(31), 33-7.

Upton, D., Upton, P. 2006. Development of an evidence-based practice questionnaire for nurses. *Journal of Advanced Nursing* 54(4), 454-458.

Waldenstrom, U. 1998. Continuity of carer and satisfaction. *Midwifery* 14, 207-13.

Walker, J.M., Hall, S., Thomas, M. 1995. The experience of labour: a perspective from those receiving care in a midwifery-led unit. *Midwifery* 11(3), 120-9.

Wax, J., Cartin, A., Pinette, M., Blackstone, J. 2004. Patient choice caesarean: an evidence based view. *Obstet Gynecol Surv* 59, 601-616.

Table 1: Themes and subthemes of the Birth Satisfaction Scale (BSS)

(1) Quality of care provision	(1a) Home assessment (1b) Birth environment (1c) Sufficient support (1d) Relationships with health care professionals
(2) Personal attributes	(2a) Ability to cope during labour (2b) Feeling in control (2c) Preparation for childbirth (2d) Relationship with baby
(3) Stress experienced during labour	(3a) Distress experienced during labour (3b) Obstetric injuries (3c) Perception of having received sufficient medical care (3d) Receipt of an obstetric intervention (3e) Pain experienced (3f) Long labour (3e) Health of baby

Table 2: Hollins Martin Birth Satisfaction Scale (BSS) (Hollins Martin and Fleming, 2011)

- (1) I coped well during my birth.
- (2) The delivery room staff encouraged me to make decisions about how I wanted my birth to progress.
- (3) I was well prepared for my labour, i.e., read a lot of literature and/or attended parenthood education classes.
- (4) I found giving birth a distressing experience.
- (5) I came through childbirth virtually unscathed.
- (6) I gave birth to a healthy normal baby.
- (7) During labour I received outstanding medical care.
- (8) I received a lot of medical intervention, i.e., induction, forceps, section etc.
- (9) I had a swift and speedy labour.
- (10) I felt well supported by my partner during labour and birth.
- (11) I was encouraged to hold my baby for a substantial amount of time after birth.
- (12) My birth experience was considerably different to what I intended.
- (13) I had the same midwife throughout the entire process of labour and delivery.
- (14) I felt that the delivery room was unthreatening and comfortable.
- (15) I felt very anxious during my labour and birth.
- (16) I felt out of control during my birth experience.
- (17) I felt it was better not to know in advance about the processes of giving birth.
- (18) I was not distressed at all during labour.
- (19) I felt mutilated by my birth experience.
- (20) My baby was avoidably hurt during birth.
- (21) The staff provided me with insufficient medical care during my birth.
- (22) I had a natural labour, i.e., minimal medical intervention.
- (23) I thought my labour was excessively long.
- (24) I felt well supported by staff during my labour and birth.
- (25) I was separated from my baby for a considerable period of time after my birth.
- (26) My birth proceeded as I planned it.
- (27) The staff communicated well with me during labour.
- (28) The delivery room was clean and hygienic.
- (29) Giving birth was incredibly painful.
- (30) Labour was not as painful as I imagined.

Table 3. Characteristics of participants as a function of birthing unit type, delivery type and method of pain relief (N=228).

	Primigravida (N=110)	Multigravida (N=118)
<i>Birthing unit</i>		
Delivery suite	68 (62%)	82 (69%)
Birthing unit	11 (10%)	14 (12%)
Home	0 (0%)	2 (2%)
Theatre	30 (27%)	20 (17%)
Unknown	1 (1%)	0 (0%)
<i>Type of delivery</i>		
Normal vaginal	51 (46%)	91 (77%)
Forceps	24 (22%)	6 (5%)
Ventouse	12 (11%)	1 (1%)
Pre-organised CS*	1 (1%)	12 (10%)
Emergency CS	21 (19%)	4 (3%)
Breech	0 (0%)	0 (0%)
Pool	1 (1%)	4 (3%)
<i>Pain relief</i>		
Epidural	65 (59%)	35 (30%)
Opiates	15 (14%)	13 (11%)
Entonox	25 (23%)	63 (53%)
Alternative methods**	2 (2%)	1 (1%)
None	2 (2%)	6 (5%)
Unknown	1 (1%)	0 (0%)

*Caesarean Section ** e.g., reflexology, acupuncture, homeopathy

Table 4. Individual item distributional characteristics of the Birth Satisfaction Scale (BSS) (N=228).

BSS item	Mean	SD	Skew	Kurtosis
BSS 1	4.13	0.68	-0.84	1.69
BSS 2	4.07	0.84	-0.95	1.38
BSS 3	4.07	0.84	-0.74	0.31
BSS 4	3.48	1.12	-0.45	-0.63
BSS 5	3.40	1.13	-0.41	-0.82
BSS 6	4.83	0.44	-3.32	14.62
BSS 7	4.38	0.84	-1.65	3.10
BSS 8	3.41	1.47	-0.35	-1.34
BSS 9	3.44	1.27	-0.45	-0.90
BSS 10	4.72	0.58	-2.67	9.74
BSS 11	4.43	0.85	-1.57	2.01
BSS 12	2.81	1.29	0.05	-1.12
BSS 13	3.58	1.35	-0.48	-1.18
BSS 14	4.37	0.65	-1.32	4.64
BSS 15	3.30	1.22	-0.37	-0.95
BSS 16	3.61	1.10	-0.67	-0.18
BSS 17	3.76	1.10	-0.64	-0.19
BSS 18	2.95	1.14	0.15	-0.85
BSS 19	4.23	0.86	-0.80	-0.30
BSS 20	4.33	1.06	-1.69	2.01
BSS 21	4.24	1.09	-1.58	1.80
BSS 22	3.50	1.47	-0.52	-1.23
BSS 23	3.86	1.11	-0.94	0.27
BSS 24	4.47	0.76	-1.95	5.27
BSS 25	4.54	0.83	-2.21	5.10
BSS 26	3.03	1.25	-0.11	-0.95
BSS 27	4.52	0.67	-1.70	4.25
BSS 28	4.69	0.52	-1.41	1.04
BSS 29	2.76	1.22	0.19	-0.85
BSS 30	2.92	1.22	0.09	-0.90

Table 5. The 10 Birth Satisfaction Scale items comprising three-factor best-fit structural and measurement model and associated domains.

Note: Original item number from the 30-item BSS in parentheses.

Item	Question	Domain
1	(5) I came through childbirth virtually unscathed.	Stress
2	(23) I thought my labour was excessively long.	Stress
3	(2) The delivery room staff encouraged me to make decisions about how I wanted my birth to progress.	Quality
4	(15) I felt very anxious during my labour and birth	Attributes
5	(24) I felt well supported by staff during my labour and birth.	Quality
6	(27) The staff communicated well with me during labour.	Quality
7	(4) I found giving birth a distressing experience.	Stress
8	(16) I felt out of control during my birth experience.	Attributes
9	(18) I was not distressed at all during labour.	Stress
10	(28) The delivery room was clean and hygienic.	Quality

Table 6. Correlations between Birth Satisfaction Scale-Revised (BSS-R) total scores and sub-scale scores. All correlations statistically significant at $p < 0.01$ (N=228).

Scale	BSS-R total	BSS-SL	BSS-WA	BSS-QC
BSS-R total		0.86	0.80	0.63
BSS-SL			0.57	0.26
BSS-WA				0.35
BSS-QC				

Key to sub-scales: *quality of care provision* (BSS-QC), *women's personal attributes* (BSS-WA), and *stress experienced during labour* (BSS-SL)

Table 7. Mean Birth Satisfaction Scale-Revised (BSS-R) total scores and sub-scale scores as a function of delivery type. Standard deviations in parentheses (N=223).

Variable	Normal Delivery	Non-normal delivery
BSS-R total	29.19 (5.86)	26.51 (5.13)
BSS-SL	10.39 (3.18)	8.30 (3.05)
BSS-WA	5.11 (2.05)	4.46 (1.82)
BSS-QC	13.68 (2.17)	13.75 (2.05)

Key to sub-scales: *quality of care provision* (BSS-QC), *women's personal attributes* (BSS-WA), and *stress experienced during labour* (BSS-SL)

Table 8: Valid and reliable Birth Satisfaction scale Revised (BSS-R)

If you would like to use the BSS-R please contact the first author at:
c.j.hollins-martin@salford.ac.uk

(1) I came through childbirth virtually unscathed.

Strongly Agree Agree Neither Agree or Disagree Disagree Strongly Disagree

Comments _____

(2) I thought my labour was excessively long.

Strongly Agree Agree Neither Agree or Disagree Disagree Strongly Disagree

Comments _____

(3) The delivery room staff encouraged me to make decisions about how I wanted my birth to progress.

Strongly Agree Agree Neither Agree or Disagree Disagree Strongly Disagree

Comments _____

(4) I felt very anxious during my labour and birth.

Strongly Agree Agree Neither Agree or Disagree Disagree Strongly Disagree

Comments _____

(5) I felt well supported by staff during my labour and birth.

Strongly Agree Agree Neither Agree or Disagree Disagree Strongly Disagree

Comments _____

(6) The staff communicated well with me during labour.

Strongly Agree Agree Neither Agree or Disagree Disagree Strongly Disagree

Comments _____

(7) I found giving birth a distressing experience.

Strongly Agree Agree Neither Agree or Disagree Disagree Strongly Disagree

Comments _____

(8) I felt out of control during my birth experience.

Strongly Agree Agree Neither Agree or Disagree Disagree Strongly Disagree

Comments _____

(9) I was not distressed at all during labour.

Strongly Agree Agree Neither Agree or Disagree Disagree Strongly Disagree

Comments _____

(10) The delivery room was clean and hygienic.

Strongly Agree Agree Neither Agree or Disagree Disagree Strongly Disagree

Comments _____

(Q1) I coped well during my birth.					
	Strongly Agree	Agree	Neither Agree or Disagree	Disagree	Strongly Disagree
Scores	5	4	3	2	1
Comments	_____				

Figure 1. Example question from the Birth Satisfaction Scale showing the specific question (Q1), Likert rating and individual question scoring, and additional comments.

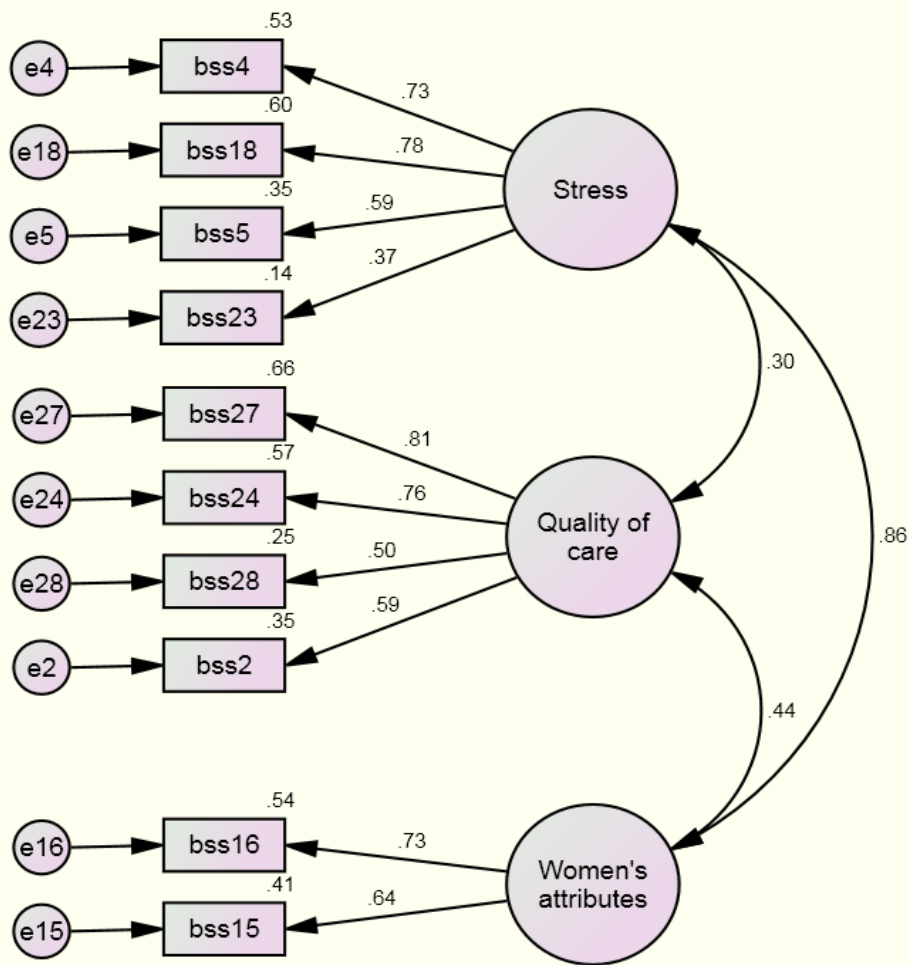


Figure 2. Measurement model of the ten-item Revised version of the Birth Satisfaction Scale (BSS-R) (estimates are standardised)

Note: Details of the BSS-R items in Figure 2. can be found with reference to Table 2. The arrows leading from the factor (circle) to the BSS item (box) represents the standardised regression coefficient weight of the factor on the indicator (BSS item). The value by each box (BSS item) indicates the proportion of variance of the item explained by the factor. As an example in the case of BSS item 4, this would be 53% which would also mean that 47% of variance of this item was unexplained or error 'e' variance. The double-headed arrow between the factors represents the covariance between factors.

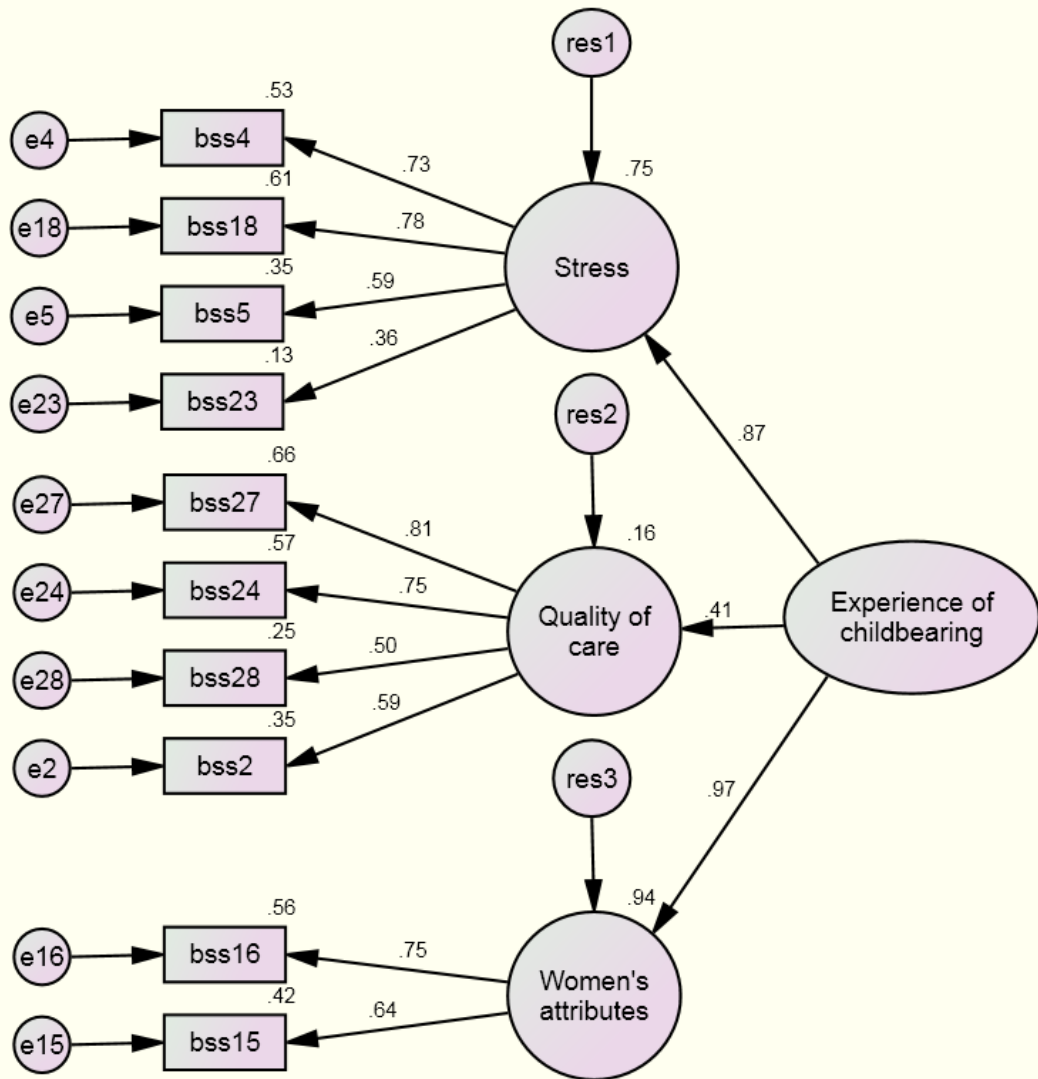


Figure 3. Second order model of the Birth Satisfaction Scale (BSS-R) with an experience of childbearing factor underpinning stress, quality and attributes factors (estimates are standardised)