Development and validation of the Australian version of the Birth Satisfaction Scale-Revised (BSS-R)

Elaine Jefford[1]

Caroline J. Hollins Martin[2]

and

Colin R. Martin[3]

[1]Research Lead - Midwifery, School of Health & Human Sciences, Southern Cross University, Coffs Harbour, New South Wales 2450, Australia.

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

[3*]Professor of Mental Health, Faculty of Society and Health, Buckinghamshire New University, Uxbridge, UK, UB8 1NA.

*Corresponding author

Address for correspondence: Professor Colin R Martin, Room 2.11, Faculty of Health and Society, Buckinghamshire New University, Uxbridge Campus 106 Oxford Road, Uxbridge, Middlesex, UB8 1NA, UK. Tel: 01494 522141 Extension 2349; Fax: 01494 603179; Email: colin.martin@bucks.ac.uk

Development and validation of the Australian version of the Birth Satisfaction Scale-Revised (BSS-R)

Abstract

Objective and background: The 10-item Birth Satisfaction Scale-Revised (BSS-R) has recently been endorsed by international expert consensus for global use as the birth satisfaction outcome measure of choice. English-language versions of the tool include validated UK and US versions, however the instrument has not to date, been contextualised and validated in an Australian English-language version. The current investigation sought to develop and validate an English-language version of the tool for use within the Australian context.

Methods: A two-stage study. Following review and modification by expert panel, the Australian BSS-R (A-BSS-R) was (Stage 1.) evaluated for factor structure, internal consistency, known-groups discriminant validity and divergent validity. Stage 2. directly compared the A-BSS-R dataset with the original UK dataset to determine the invariance characteristics of the new instrument. Participants were a purposive sample of Australian postnatal women (n=198).

Results: The A-BSS-R offered a good fit to data consistent with the BSS-R tri-dimensional measurement model and was found to be conceptually and measurement equivalent to the UK version. The A-BSS-R demonstrated excellent known-groups discriminant validity, generally good divergent validity and overall good internal consistency.

Conclusion: The A-BSS-R represents a robust and valid measure of the birth satisfaction concept suitable for use within Australia and appropriate for application to International comparative studies.

Key Words: Birth Satisfaction Scale-Revised (BSS-R), Australian, validation, maternity care, measurement invariance, midwives, obstetricians

Introduction

The construct of birth satisfaction has received much recent attention, this being unsurprising given that it is a complex concept, encompassing emotional and psychological attributes at a time when the woman herself is experiencing significant physical challenges as she transitions between late labour, birth and the postnatal period (Sawyer et al., 2013). Defining birth satisfaction represents a challenge since it represents a subjective appraisal of maternity care provision influenced by the woman's attitudes and beliefs and encapsulated within a context of uniquely individualised experience (Gibbens & Thomson, 2001; C.J. Hollins Martin & Fleming, 2011; C. R. Martin et al., 2016a; Sawyer et al., 2013).

Measurement of birth satisfaction facilitates an understanding of birth experience from the perspective of the woman herself; this is important and entirely consistent with innovations in health policy that place the woman, her well-being and her birth experience at the centre of care (Department of Health, 1993). Further, examining birth satisfaction enables understanding of relationships with deleterious maternal outcomes, such as post-partum post-traumatic stress disorder (Dale-Hewitt, Slade, Wright, Cree, & Tully, 2012; Kouros, 2013; Sorenson & Tschetter, 2010), postnatal depression (Anding, Rohrle, Grieshop, Schucking, & Christiansen, 2016; Razurel & Kaiser, 2015; Webster et al., 2006; Yelland, Sutherland, & Brown, 2010), and poor mother-infant/mother-child relationships (Iles, Slade, & Spiby, 2011; McDonald, Slade, Spiby, & Iles, 2011).

The criticality of the birth satisfaction concept to maternal experience and outcome has led to the development of a range of measures to assess the concept (Goodman, Mackey, & Tavakoli, 2004; Harvey, Rach, Stainton, Jarrell, & Brant, 2002; Redshaw & Martin, 2009), however, a number of shortcomings related to these existing measures have been highlighted (Sawyer et al., 2013). Sawyer and colleagues (2013) highlight that measures and tools that have been used to assess satisfaction with the childbirth experience have invariably not been based on cogent theoretical accounts of satisfaction and in many instances were specific in application, for example, those women who have had a Caesarean section. A recently developed measure that has gained considerable traction internationally is the Birth Satisfaction Scale-Revised (BSS-R; C. J. Hollins Martin & Martin, 2014). Developed from the thematically derived 30-item Birth Satisfaction Scale (C.J. Hollins Martin & Fleming, 2011), the 10-item BSS-R assesses three domains of (i) Stress experienced during child-bearing, (ii), Women's attributes, and (iii), Quality of care. Contrasting with Sawyer et al.'s (2013) critique of existing measures, it is important to note that the Birth Satisfaction Scale was developed from an exhaustive thematic review of the existing literature, grounded in theoretical authenticity within the field of satisfaction and the measure developed for inclusive use within the birth context (C.J. Hollins Martin & Fleming, 2011). Consequently, rather than being specific to a particular clinical circumstance of birth experience, the measure was developed to be suitable for use in all childbearing women. Being a direct derivative measure of the birth satisfaction scale, the BSS-R shares and is underpinned by the same conceptual and theoretical framework (C. J. Hollins Martin & Martin, 2014). Thus consistent with the theoretical and empirical architecture of the birth satisfaction scale, the 'quality of care' domain for example, is underpinned by the same extensive

applied literature condensed to sub-themes during the original thematic review, such as *birth environment, support* and *relationship with health professionals*. The instrument developers reported excellent psychometric properties of this multi-dimensional tool while the inherent domains of the scale remained conceptually consistent with the thematically-derived longer measure (C. J. Hollins Martin & Martin, 2014). The BSS-R has been validated and translated into Greek (Vardavaki, Hollins Martin, & Martin, 2015) and a United States (Barbosa-Leiker, Fleming, Hollins Martin, & Martin, 2015) version of the tool, with these studies suggesting excellent validity and reliability. The US version of the BSS-R has recently been used in a large-scale US study (Fleming et al., 2016) and a psychometric evaluation of the data from that study corroborated not only the excellent psychometric properties of the US version, but also its measurement robustness when comparing discrete sub-groups within the sample (C. R. Martin et al., 2016a). A comparison of UK and Greek BSS-R datasets has revealed the instrument to be both conceptually and statistically comparable (C.R. Martin, Vardavaki, & Hollins Martin, 2016b), thus emphasising the potential utility of the measure for meaningful international comparative studies that focus on birth experience and outcomes. Consistent with this observation, Burduli and colleagues (2017) have found UK and US BSS-R validation datasets to be broadly equivalent (Burduli, Barbosa-Leiker, Fleming, Hollins Martin, & Martin, 2017). The robust psychometric properties of the BSS-R reported to date and the consistency of these findings between translated versions has been salient evidence in the recent adoption of the measure as the recommended measure of birth satisfaction for global health outcome measurement use in the recently published *Pregnancy and Childbirth Standard Set* (International Consortium for Health Outcomes Measurement (ICHOM), 2016). Since the publication of the ICHOM

guidance, there have been over 40 requests to translate or use the BSS-R in birth outcome studies throughout the world[1].

Australia represents a developed country with a highly advanced health economy that incorporates features of both UK and US service delivery models (Jackson, Dimitropoulos, Madden, & Gillett, 2015; Ragupathy, Aaltonen, Tordoff, Norris, & Reith, 2012). A pertinent feature of Australian maternity care services are that they remain highly medicalised (Benoit, Zadoroznyj, Hallgrimsdottir, Treloar, & Taylor, 2010; Weaver, Clark, & Vernon, 2005) despite the strong movement led by government reforms towards midwifery-led models (Quinn, Noble, Seale, & Ward, 2013; Tracy et al., 2013; Tracy et al., 2014). However, the movement toward woman-centred care and choice has gathered some pace in Australia (Catling & Homer, 2016) and similar to both the UK and the US (Fleming et al., 2016; C. J. Hollins Martin & Martin, 2014) an important component of both service transition and service evaluation is women's satisfaction with their birth experience (K. Clark, Beatty, & Reibel, 2016; Lewis, Hauck, Ronchi, Crichton, & Waller, 2016). As Australia transitions from the medical paradigm characteristic of the US towards the woman-centred models of care more typical of the UK, accurate assessment of birth satisfaction is of fundamental importance.

---

[1] Communication with instrument developers.

The BSS-R would appear to be ideally suited as the measure of choice within this context, however an Australian version of the BSS-R has to date, yet to be developed and psychometrically evaluated. It is noteworthy that there are subtle differences between the UK and US versions of the tool, specifically in relation to wording of question 1. '*I came through childbirth virtually unscathed*' (UK) and '*I came through childbirth virtually unharmed*' (US). Barbosa-Leiker et al. (2015) and C. R. Martin et al. (2016a) highlight the importance of using the most appropriately specified and worded items for the population where the tool is to be used. Though predominantly an English-speaking country, Australia has significant geo-political links and maternity service delivery similarities with the US and historical cultural affiliations with the UK. It is therefore not readily apparent whether either the US or UK version of the BSS-R could be used within an Australian context without either modification or psychometric evaluation. A recent study by Burduli et al. (2017) illustrates the desirability for evaluating the measurement characteristics in the population that the tool is to be used in. Burduli and colleagues (2017) found that UK and US versions of the BSS-R were equivalent within their respective populations. Burduli et al. (2017) noted that though the measurement characteristics of the BSS-R were equivalent, it was noteworthy that US women had comparatively lower levels of birth satisfaction. Confidence in the robustness of this finding (differences in birth satisfaction) is largely contingent on the equivalence of the measurement properties of the tool confirmed in Burduli et al.'s (2017) study. Therefore two fundamental issues should be addressed in the development of an Australian version of the tool. Firstly, that the measure is contextually appropriate for the population and secondly, it is equivalent in terms of measurement to other versions of the BSS-R in order that results from studies can be meaningfully compared.

The current investigation sought to develop the Australian version of the BSS-R (A-BSS-R) and evaluate the psychometric properties of the tool. The objectives of the current study are to:

1. Demonstrate the replicability of the tri-dimensional measurement model of the BSS-R to the A-BSS-R.

2. Evaluate the internal consistency of the Quality of Care (QC), Women's Attributes (WA), and Stress Experienced during Child-bearing (SE) sub-scales anticipated to comprise the A-BSS-R.

3. Evaluate the known-groups discriminant validity of the A-BSS-R.

4. Evaluate the convergent validity of the A-BSS-R.

5. Evaluate the equivalence between the A-BSS-R and the original BSS-R.

Method

A two-stage design was used. Stage 1 (Objectives 1-4) involved a cross-sectional survey design utilising purposive sampling. Inclusion criteria included women of 18 years of age and over and current participation in the Continuity of Care Experience (CoCE; Australian Nursing and Midwifery Accreditation Council, 2014; Ebert, Tierney, & Jones, 2016; Tierney, Sweet, Houston, & Ebert, 2016) programme. Every undergraduate and postgraduate midwifery program leading to registration as a midwife has to encompass Continuity of Care Experience (CoCE). A midwifery student must 'follow' at least 10 women across their unique childbearing journey. The aim is to immerse midwifery students within the very essence of midwifery and being 'with woman'. Student midwives recruit pregnant women in to the CoCE programme in a variety of ways, for example pamphlets/flyers at antenatal clinics or word of mouth. CoCE programmes provide students with an opportunity to build a trusting partnership relationship within a woman-centred ethos through their pregnancy, labour and early parenting time. Women enrolled in the CoCE programme via communication with the Professional Experience Unit were invited to participate in an online survey following completion of the CoCE programme having been given pre-recruitment information that they would be invited to voluntarily participate in a survey about their birth experience on CoCE programme completion. Following participant review of the participant information document and consent being given participants were able to access the survey. Women were invited to complete the A-BSS-R within six weeks of birth.

Stage 2 (Objective 5) involved a direct psychometric comparison of the dataset from the current study with that of the original UK dataset (Hollins Martin and Martin, 2014). The stage 2 component of the study is differentiated from stage 1 because of the use of an existing BSS-R dataset to facilitate comparisons with A-BSS-R data collected in the current investigation.

Ethical approval

Ethical approval for the study was granted by ███████████████████████ ███████████, Australia.

Participants

Participants were purposively sampled postnatal women who were currently taking part in the CoCE programme.

Measures

The BSS-R (C. J. Hollins Martin & Martin, 2014) comprises ten items scored on a five-point Likert type scale with responses ranging from (i.) strongly agree, (ii.) agree, (iii.) neither agree or disagree, (iv.) disagree, (v.) strongly disagree. Six of the BSS-R items are scored from 4 (strongly agree) to 0 (strongly disagree). The remaining four BSS-R items are reverse-scored thus 0 (strongly agree) to 4 (strongly disagree). The *stress experienced during child-bearing* and *quality of care* sub-scales each comprise four items. The *women's attributes* sub-scale comprises two items. BSS-R sub-scale scores are simply calculated as the sum score of the items representing that specific BSS-R domain. A total BSS-R score can also be calculated (range 0-40). Higher scores correspond to comparatively higher birth satisfaction across all sub-scales and the total score. The BSS-R has been found to be psychometrically robust, valid and reliable, with replication of the underlying three-factor measurement model in both English-language (Barbosa-Leiker et al., 2015; Burduli et al., 2017; C. R. Martin et al., 2016a) and translated (C.R. Martin et al., 2016b; Vardavaki et al., 2015) versions.

Development of the Australian version of the BSS-R: Item review

The items of the existing English-language versions of the BSS-R were reviewed by an expert panel of Australian psychologists (N=2), Australian midwives (N=5) and the UK developers of the BSS-R to ensure contextual appropriateness of items between versions. Following review, BSS-R item 3. *'The delivery room staff encouraged me to make decisions about how I wanted my birth to progress'* was modified slightly to *'The birthing room staff encouraged me to make decisions about how I wanted my birth to progress'.* The panel felt the US version of BSS-R item 1. *'I came through childbirth virtually unharmed'* was more contextually appropriate for Australian women than the original UK version of this item *'unscathed'.* Further, a convenience sample of student midwives (N=49) were asked which version of BSS-R item 1 they believed was more appropriate for communicating with Australian women, and all indicated the US version of this item. No other changes or modifications were made to the A-BSS-R. The scoring system is identical to that of the UK and US versions, thus comparatively higher scores indicate greater satisfaction with respect to the specific BSS-R sub-scale domain.

Statistical analysis (stage 1)

Objective 1 was addressed using Confirmatory Factor Analysis (CFA)(Brown, 2015). CFA is a structural equation modelling (SEM) approach to model evaluation (Byrne, 2010). CFA and SEM are parametric techniques and thus bound by the accompanying underlying assumptions regarding data normality (Brown, 2015; Byrne, 2010). Scrutiny of data for excessive skew and kurtosis and removal of multivariate outliers is thus necessary to determine any potential violation of the underlying parametric assumptions that could lead to an incorrect interpretation of the statistical analysis (P. Kline, 2000). Two CFA models were evaluated, which were the tri-dimensional measurement model of the BSS-R comprising correlated factors of SE, WA and QC specified by Hollins Martin and Martin (2014) and a single factor version of this model (correlations between factors set to 1). A hierarchical model which specified experience of childbirth as a second-order domain, was not evaluated given the recent findings of Burduli et al. (2017) of the psychometric superiority of the tri-dimensional model of the BSS-R in English-language versions of the instrument. Assuming multivariate normality, a maximum-likelihood estimation approach was taken to model evaluation (Brown, 2015; R. B. Kline, 2011). The two models were evaluated by consensus of a range of model fit indices (Bentler & Bonett, 1980), these being the comparative fit index (CFI; Bentler, 1990), the root mean squared error of approximation (RMSEA; Steiger & Lind, 1980), the square root mean residual (SRMR; Hu & Bentler, 1999). The χ^2 statistic may be used to evaluate model fit, with a non-significant p value indicating that the major proportion of variance within the data is explained by the model. Importantly, χ^2 is influenced by both sample size and data variation, thus model evaluation is seldom based on χ^2

and invariably based on the model fit indices highlighted above (Byrne, 2010; Hooper, Coughlan, & Mullen, 2008; Vardavaki et al., 2015).

Divergent validity

Divergent validity was evaluated by correlating A-BSS-R sub-scale scores with the number of weeks pregnant at birth (C. J. Hollins Martin & Martin, 2014). No statistically significant correlation between A-BSS-R sub-scale scores and this clinical indices is predicted.

Known-groups discriminant validity

Known-groups discriminant validity was evaluated using the approach of Vardavaki et al. (2015). Stratification of the dataset was undertaken following data collection by categorising by delivery type, this being either unassisted vaginal delivery or assisted vaginal delivery. It was predicted that A-BSS-R total, and A-BSS-R SE and A-BSS-R WA sub-scale scores would be significantly higher in those women who had an unassisted vaginal delivery compared to those who either had an instrument, ventouse, Caesarean Section or emergency Caesarean Section. It was predicted that there would be no statistically significant differences between groups on the A-BSS-R QC sub-scale.

Internal consistency

The internal consistency of the A-BSS-R sub-scales and total score was evaluated using Cronbach's coefficient alpha (Cronbach, 1951). A Cronbach's alpha of 0.70 or above is generally recognised as a threshold for acceptable internal consistency (P. Kline, 2000). Cronbach's alpha represents a lower-bound internal consistency estimation (Sijtsma, 2009) and is deflated by low item N within a scale (Cortina, 1993; Schmitt, 1996). Given that the A-BSS-R WA sub-scale comprises just two items, the inter-item correlation (Pearson's *r*) between these items was scrutinised since L. A. Clark and Watson (1995) suggest that inter-item correlations between 0.15-0.50 indicate acceptable reliability within a scale. This approach is consistent with Eisinga, Grotenhuis, and Pelzer (2013) who suggest this statistical treatment may be preferable to Cronbach's alpha for evaluating a two-item scale.

Measurement equivalence (stage 2)

Objective 5 was addressed by a direct comparison of the A-BSS-R dataset and its tri-dimensional model specification with the original UK BSS-R dataset, again specified within the same tri-dimensional model. The procedure to determine measurement invariance requires the application of increasingly restrictive versions of the underlying model (Brown, 2015; C. R. Martin et al., 2016a). An initial diagnostic approach prior to the measurement invariance procedure being conducted is to ensure a theoretically plausible and statistically satisfactory fit to the specified model within the dataset (Brown, 2015). The original UK BSS-R validation dataset satisfied this criteria (C. J. Hollins Martin & Martin, 2014), however, measurement invariance testing within the current study will only be conducted in the event of the A-BSS-R demonstrating a satisfactory fit to data for the tri-dimensional model, such evidence being provided by the findings from the CFA. Following the approach of C.R. Martin et al. (2016b) and contingent on satisfactory A-BSS-R model fit, an omnibus model fit evaluation of the pooled dataset (UK and Australian) will be conducted in the absence of constraints (using the same fit measures and associated as CFA outlined earlier). A configural invariance model will then be evaluated to determine equivalence of the factor model and pattern of loadings between groups (A-BSS-R, UK BSS-R). A satisfactory configural invariance model enables a more restrictive metric invariance model to be evaluated where item-factor loadings are constrained to be the same across groups. R. B. Kline (2011) highlights that metric invariance is required to demonstrate equivalence of meaning of the underlying constructs specified within the measurement model. Observation of metric invariance would enable further constraints to be made in the form of evaluation of scalar invariance in which the item intercepts are constrained to be equal between groups. Evidence of

non-invariance in an item at the scalar level is indicative of group differences on the mean of that item even within the context of having comparable values on the factor related to the item itself. It is conceivable (at each level of invariance evaluation) that some items will not be invariant between groups whereas others will be. This represents a context known as partial invariance (Byrne, 2010) and typically the process of measurement invariance stops at the level where partial invariance is found. Partial invariance has different implications depending on the level of invariance testing in which it is observed. Thus full metric invariance is considered important to confirm conceptual equivalence of meaning and measurement between the two groups (Vandenberg & Lance, 2000), and thus the ability to use the measure within both the population of interest and to make meaningful comparisons between other groups evaluated by other validated versions of the measure. Evidence of partial invariance at the scalar level, in contrasts, indicates that though the measure is equivalent at the metric level, an invariant intercept (item mean) represents a difference in mean score (though the item-factor loading is equivalent) between groups, which is both meaningful and essentially a true score difference rather than an indication of measurement error (C.R. Martin et al., 2016b; Millsap, 1998). Non-invariant items are identified by evidence of a significant difference between models (invariant/non-invariant), as evidenced by a difference in CFI of >0.01 (Cheung & Rensvold, 2002). Statistical analysis was conducted using the R programming language (R Core Team, 2017) and the R packages Lavaan (Rosseel, 2012), Effsize (Torchiano, 2015), Cocor (Diedenhofen & Musch, 2015), and Cocron (Diedenhofen & Musch, 2016).

Results

Participants

Two-hundred and seventeen participants consented to take part in the study, with 10 participants having significant missing data (>5%) and so their data was excluded from the dataset. Following this, the dataset was screened (N=207) for multivariate outliers within the A-BSS-R data stream by reference to Mahalanobis distances, and N=9 were found and excluded based on a distance from the centroid estimation of χ^2 > 34.53. The final number of participants for which data was complete and multivariate normal was thus N=198. The mean age of participants was 29.66 (SD 5.61) years with a range of 18-44. The mean duration of pregnancy was 39.78 (SD 1.42) weeks. Seventy-nine (40%) women were primigravidas. The mean scores and distributional characteristics of the A-BSS-R thematically derived sub-scales and total score are summarised in Table 1.

TABLE 1. ABOUT HERE

Individual item scores for the A-BSS-R and associated distributional characteristics are summarised in Table 2. No evidence of excessive skew or kurtosis was observed.

TABLE 2. ABOUT HERE

The A-BSS-R total score was significantly correlated with A-BSS-R SE, WA, and QE sub-scales, $r = 0.89$, $p <0.001$, $r = 0.83$, $p <0.001$ and $r = 0.60$, $p <0.001$, respectively. The A-BSS-R SE sub-scale was positively and significantly correlated with the A-BSS-R WA ($r = 0.67$, $p <0.001$) and A-BSS-R QC ($r = 0.23$, $p <0.001$) sub-scales. Finally, the A-BSS-R WA sub-scale was positively and significantly correlated with the A-BSS-R QC ($r = 0.32$, $p <0.001$). Comparisons between the above correlation r values and those reported by Hollins Martin and Martin (2014) were conducted using Fisher's Z transformation (Fisher, 1925) and Zou's confidence interval (Zou, 2007) and adopting the procedure of Diedenhofen and Musch (2015) and summarised in Table 3. No statistically significant differences between A-BSS-R and UK BSS-R total or sub-scale correlational values were observed.

TABLE 3. ABOUT HERE

Confirmatory factor analysis

The three-factor measurement model of the BSS-R offered a good fit to the data ($\chi^2_{(df=32)} = 71.19$, $p <0.01$, CFI = 0.95, RMSEA = 0.08, SRMR = 0.06). The single-factor model offered a comparatively poorer fit to the data ($\chi^2_{(df=35)} = 327.88$, $p <0.01$, CFI = 0.60, RMSEA = 0.21, SRMR = 0.14).

Divergent validity

No significant correlations were found between the A-BSS-R total score, A-BSS-R SE sub-scale scores and A-BSS-R QC sub-scale scores and the number of weeks gestation, $r = 0.05$, $p = 0.51$, $r = -0.06$, $p = 0.41$, $r = 0.07$, $p = 0.33$ respectively. Against prediction, the A-BSS-R WA sub-scale was observed to be significantly and positively correlated with the number of weeks of gestation, $r = 0.17$, $p = 0.01$.

Known-groups discriminant validity

The results of the known-groups discriminant validity analysis are summarised in Table 4. These findings are consistent with predictions, with those experiencing an unassisted vaginal delivery having significantly higher A-BSS-R SE and A-BSS-R WA sub-scale scores, and A-BSS-R total scores compared to those experiencing an assisted/operative delivery.

TABLE 4. ABOUT HERE

Internal consistency

Internal consistency analysis revealed acceptable Cronbach alphas for the A-BSS-R SE, and A-BSS-R QC, and A-BSS-R total score of 0.74, 0.81 and 0.81 respectively. Cronbach alpha of the A-BSS-R WA sub-scale was 0.66. A comparison between Cronbach alpha of this dataset with those reported by Hollins Martin and Martin (2014) revealed no statistically significant differences at sub-scale and total score level (Table 5.). The inter-item correlation between the two A-BSS-R WA sub-scale items was $r = 0.49$, $p < 0.001$.

TABLE 5. ABOUT HERE

Invariance evaluation

The results of the invariance testing are summarised in Table 6. The overall (pooled data) model (Model 1) provided an excellent fit to the data, which facilitated the invariance testing to proceed stage-wise. The configural model (Model 2.) was found to offer an excellent fit to the data across all model fit parameters. Constraining item-factor loadings equal between the Australian and UK datasets, resulted in a model (Model 3) that was not significantly different (ΔCFI ≤1) to the configural model. Scalar invariance was then evaluated and revealed a significant difference (ΔCFI >1) between this model (Model 4) and the metric model (Model 3) indicating non-invariance. Examination of modification indices suggested that the constraint between datasets for A-BSS-R item 1. *'I came through childbirth virtually unscathed'* be freely estimated and this model resulted in an improved but still significantly different model (Model 5) to the metric invariance model (Model 3) (ΔCFI >1). Additional scrutiny of modification indices suggested A-BSS-R item 3 *'The birthing room staff encouraged me to make decisions about how I wanted my birth to progress'* equality constraints be relaxed between datasets, and, in addition to A-BSS-R item 1, Model 6 was shown not to be significantly different to Model 3 (ΔCFI ≤1), thus demonstrating partial scalar invariance.

TABLE 6. ABOUT HERE

Discussion

The findings from this validation study of the Australian version of the BSS-R are broadly consistent with those of the original UK version of the instrument (C. J. Hollins Martin & Martin, 2014). Moreover, the current investigation reveals additional salient inferential statistical detail regarding the comparability of the measure to the original UK version in terms of internal consistency and sub-scale correlations. Consistent with prediction, direction, and the work of C. J. Hollins Martin and Martin (2014), the known-group discriminant validity testing revealed statistically significant differences as a function of birth type on A-BSS-R SE and A-BSS-R WA sub-scales and the total A-BSS-R score. No statistically significant difference, as anticipated, was observed as a function of birth type on the A-BSS-R QC sub-scale. This finding, though consistent with C. J. Hollins Martin and Martin (2014), is inconsistent with the Greek version of the measure where a significant difference on this sub-scale was observed (Vardavaki et al., 2015). An explanatory account suggested by Vardavaki et al. (2015), was that the significant impact of birth type on this sub-scale may have been consequential of implicit psychological/psychosocial cues of normality associated with an unassisted vaginal birth experience specific to, and within that particular cultural context. For example, a sense of reassurance, in contrast to absolute differences in care received. However, the current findings in this sample, consistent with those of the UK validation sample, would suggest that there was no perceived difference in quality of care provided. Reflecting on these observations with this contemporary evidence, it may be that there are tangible differences in experiences of quality of care within the Greek maternity system as a function of birth type compared with UK and Australian maternity service delivery. Addressing this particular issue in depth is clearly beyond the scope of the current investigation,

however, it does serve to highlight the value of a validated and comparable model of birth satisfaction that can allow meaningful comparisons between maternity service models, health economies, cultures, and countries.

The three-factor measurement model of the BSS-R was found to offer a good fit to data. Consistent with previous investigations of the BSS-R (Barbosa-Leiker et al., 2015; C. J. Hollins Martin & Martin, 2014; Vardavaki et al., 2015), a uni-dimensional model of instrument offered a comparatively poor fit to data. This observation represents both coherence with the postulated thematic structure of the birth satisfaction construct (C.J. Hollins Martin & Fleming, 2011; C. J. Hollins Martin & Martin, 2014), and the preferential use of the three sub-scales of the instrument over the total score, particularly in the context of International comparisons of the measure (Burduli et al., 2017).

The invariance evaluation provides valuable and affirming evidence regarding the comparability of the A-BSS-R with the UK version. The observation of good configural model fit and metric invariance is unambiguous in demonstrating the equivalence of conceptual meaning and tri-dimensional structure between versions (R. B. Kline, 2011; Vandenberg & Lance, 2000). The finding that not only did the A-BSS-R satisfy the criteria for partial scalar invariance, but in this regard just two items (20%) were found to be non-invariant at the scalar invariance level of analysis, thus 80% of items are invariant. Strikingly, these two non-invariant items, A-BSS-R item 1[2]., *'I came through childbirth virtually unharmed'* and A-BSS-R item 3[3]., *'The birthing room staff encouraged me to make decisions about how I wanted my birth to progress'* were the same non-invariant items observed in the Greek version of the measure (Vardavaki et al., 2015) at the scalar invariance level of evaluation (C.R. Martin et al., 2016b). This observation not only highlights comparability in terms of another (non-English) version of the tool to the original measure, but also emphasises the same likely rationale for these invariant items, namely a true and representative difference in the mean score of these two items between versions (Brown, 2015; Millsap, 1998) within the overall context of conceptual and measurement equivalence of the measure. Examination of the relative item mean differences between the current study and that of C. J. Hollins Martin and Martin (2014) of BSS-R items 1 and 3 reveal the differences to be modest (both 0.32). It is of value to speculate on any potential differences between birthing context that may impact on these individual item scores, despite the differences being small. Potential factors could be nuanced differences in the embedded woman-centred philosophy of care between Australia (Quinn et al., 2013; Tracy et al., 2013; Tracy et al., 2014) and the UK (Renfrew et al., 2014; Sandall, Soltani, Gates, Shennan, & Devane, 2016),

differences between the ratio of primagravida to multigravida women between studies, and differences between the proportions in relation to delivery type. Given that BSS-R items 1. and 3. have been shown to be scalar non-invariant in a previous study and the current study in comparison to the original measure, a future qualitative study could be of value to determine any perceptible or meaningful difference in interpretation of item meaning and how this may relate to the possible factors highlighted above.

An unexpected finding from the divergent validity testing was the finding of a statistically significant positive correlation between A-BSS-R WA sub-scale score and the number of weeks gestation. A rationale for this observation does not immediately present itself. Nonetheless, it was unanticipated in this population and may need to be explored in future studies of the A-BSS-R to determine any persistence of this observation, and if found implications thereof.

---

[2] Note: BSS-R item 1: Greek (…….unscathed), Australian (…….unharmed).
[3] Note: BSS-R item 3: Greek (…….delivery room……), Australian (…….birthing room……).

A novel feature of the current investigation was the inferential statistical comparison of the Cronbach's alpha and Pearson's *r* correlation values with those reported in the UK validation study (C. J. Hollins Martin & Martin, 2014). As far as we are aware, this is the first study to compare these statistical indices directly in relation to the BSS-R and the finding of no statistically significant differences between the two studies on either the BSS-R total score or BSS-R sub-scales offers additional corroborative evidence supporting validity and consistency across versions of the tool. Cronbach's alpha for the A-BSS-R WA sub-scale did not reach generally accepted criteria for acceptable internal consistency (P. Kline, 2000), though there remains debate over the suitability of this statistical approach for two-item scales (Eisinga et al., 2013). Interestingly, this sub-scale was found to be acceptable according to the inter-item correlation criteria of L. A. Clark and Watson (1995). Given that the WA sub-scale is comprised of just two items, and as noted by (George & Mallery, 2003), an alpha of less than 0.70 is not necessarily unacceptable, the findings from the comparative analysis as a whole present evidence of consistency across measures.

One limitation of the current study was that we did not collect data on marital status and occupational status. We recognise that these may be important variables both to describe the population and also to examine these variables with respect to any potential influence on birth satisfaction scores. A further limitation is that we did collect details of the exact timing of questionnaire administration following birth within the six-week postpartum data collection period. This data would be useful in examining any potential correlational relationships between time from birth and birth satisfaction scores. Future studies would benefit from collecting this data where possible for the reasons described above.

Conclusion

The current investigation has sought to develop and validate the Australian version of the BSS-R. The extensive battery of psychometric evaluative approaches used in the current investigation indicate that the A-BSS-R, while modified slightly to be contextually anchored within the Australian context, is from a conceptual and measurement perspective equivalent to the original UK version. Thus comparison between scores from differing countries can be plausibly compared with confidence. A small number of issues were identified during the validation process which warrant further study, include the unexpected correlation between number of week's gestation and the A-BSS-R WA sub-scale. Also, the sample size in this current study was a realistic minimum for a validation study of this kind, and would benefit from future studies using the A-BSS-R to confirm the findings from this initial investigation using a larger participant population.

Obtaining the A-BSS-R

To request a copy of the A-BSS-R and the associated A-BSS-R scoring grid, please contact Professor Caroline Hollins Martin by email: C.HollinsMartin@napier.ac.uk

References

Anding, J. E., Rohrle, B., Grieshop, M., Schucking, B., & Christiansen, H. (2016). Couple comorbidity and correlates of postnatal depressive symptoms in mothers and fathers in the first two weeks following delivery. *Journal of Affective Disorders, 190*, 300-309. doi: 10.1016/j.jad.2015.10.033

Australian Nursing and Midwifery Nursing Accreditation Council (2014). *Midwife Accreditation Standards*. Canberra, Australia.

Barbosa-Leiker, C., Fleming, S., Hollins Martin, C. J., & Martin, C. R. (2015). Psychometric properties of the Birth Satisfaction Scale-Revised (BSS-R) for US mothers. *Journal of Reproductive and Infant Psychology, 33*(5), 504-511. doi: 10.1080/02646838.2015.1024211

Benoit, C., Zadoroznyj, M., Hallgrimsdottir, H., Treloar, A., & Taylor, K. (2010). Medical dominance and neoliberalisation in maternal care provision: the evidence from Canada and Australia. *Social Science and Medicine, 71*(3), 475-481. doi: 10.1016/j.socscimed.2010.04.005

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107*(2), 238-246

Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the evaluation of covariance structures. *Psychological Bulletin, 88*, 588-606

Brown, T. (2015). *Confirmatory Factor Analysis for Applied Research* (2nd ed.). New York: Guilford Press.

Burduli, E., Barbosa-Leiker, C., Fleming, S. E., 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*. doi: 10.1080/02646838.2017.1310374

Byrne, B. M. (2010). *Structural Equation Modeling with AMOS: Basic Concepts,*

    Applications and Programming (2nd ed.). New York: Routledge/Taylor and

    Francis Group.

Catling, C. J., & Homer, C. S. (2016). Twenty-five years since the Shearman Report:

    How far have we come? Are we there yet? *Women and Birth, 29*(1), 93-99.

    doi: 10.1016/j.wombi.2015.08.011

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating Goodness-of-Fit Indexes for

    Testing Measurement Invariance. *Structural Equation Modeling: A*

    Multidisciplinary Journal, 9(2), 233-255. doi: 10.1207/S15328007SEM0902_5

Clark, K., Beatty, S., & Reibel, T. (2016). Maternity-care: measuring women's

    perceptions. *International Journal of Health Care Quality Assurance, 29*(1),

    89-99. doi: 10.1108/IJHCQA-06-2015-0078

Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective

    scale development *Psychological Assessment, 7*(3), 309-319

Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and

    applications. *Journal of Applied Psychology, 79*, 98-104

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests.

    Psychometrika, 16(3), 297–334

Dale-Hewitt, V., Slade, P., Wright, I., Cree, M., & Tully, C. (2012). Patterns of

    attention and experiences of post-traumatic stress symptoms following

    childbirth: an experimental study. *Archives of Womens Mental Health, 15*(4),

    289-296. doi: 10.1007/s00737-012-0290-2

Department of Health. (1993). *Changing Childbirth: Report of the Expert Maternity*

    Group (Cumberlege Report). London: HMSO.

Diedenhofen, B., & Musch, J. (2015). cocor: A comprehensive solution for the

statistical comparison of correlations. *PloS One, 10*(4), e0121945. doi:

doi:10.1371/journal.pone.0121945

Diedenhofen, B., & Musch, J. (2016). cocron: A web interface and R package for the

statistical comparison of Cronbach's alpha coefficients. *International Journal of*

Internet Science, 11(1), 51–60

Ebert, L., Tierney, O., & Jones, D. (2016). Learning to be a midwife in the clinical

environment; tasks, clinical practicum hours or midwifery relationships. *Nurse*

Education in Practice, 16(1), 294-297. doi: 10.1016/j.nepr.2015.08.003

Eisinga, R., Grotenhuis, M., & Pelzer, B. (2013). The reliability of a two-item scale:

Pearson, Cronbach, or Spearman-Brown? *International Journal of Public*

Health, 58(4), 637-642. doi: 10.1007/s00038-012-0416-3

Fisher, R. A. (1925). *Statistical Methods for Research Workers*. Edinburgh, Scotland:

Oliver and Boyd.

Fleming, S. E., Donovan-Batson, C., Burduli, E., Barbosa-Leiker, C., Hollins Martin,

C. J., & Martin, C. R. (2016). Birth Satisfaction Scale/Birth Satisfaction Scale-

Revised (BSS/BSS-R): A large scale United States planned home birth and

birth centre survey. *Midwifery, 41*, 9-15. doi: 10.1016/j.midw.2016.07.008

George, D., & Mallery, P. (2003). *SPSS for Windows Step by Step: A Simple Guide*

and Reference. 11.0 update (4th ed.). Boston, MA: Allyn and Bacon.

Gibbens, J., & Thomson, A. (2001). Women's expectations and experiences of

childbirth. *Midwifery, 17*, 302-313

Goodman, P., Mackey, M. C., & Tavakoli, A. S. (2004). Factors related to childbirth

satisfaction. *Journal of Advanced Nursing, 46*(2), 212-219. doi:

10.1111/j.1365-2648.2003.02981.x

Harvey, S., Rach, D., Stainton, M., Jarrell, J., & Brant, R. (2002). Evaluation of

satisfaction with midwifery care. *Midwifery, 18*, 260-267

Hollins Martin, C. J., & Fleming, V. (2011). The birth satisfaction scale. *International

Journal of Health Care Quality Assurance, 24*(2), 124-135. doi:

10.1108/09526861111105086

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

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

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

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. (1999). Cutoff criteria for fit indexes in covariance structure

analysis: Conventional criteria versus new alternatives. *Structural Equation

Modeling, 6*, 1-55

Iles, J., Slade, P., & Spiby, H. (2011). Posttraumatic stress symptoms and

postpartum depression in couples after childbirth: the role of partner support

and attachment. *Journal of Anxiety Disorders, 25*(4), 520-530. doi:

10.1016/j.janxdis.2010.12.006

Jackson, T., Dimitropoulos, V., Madden, R., & Gillett, S. (2015). Australian diagnosis

related groups: Drivers of complexity adjustment. *Health Policy, 119*(11),

1433-1441. doi: 10.1016/j.healthpol.2015.09.011

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

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

London: Guilford Press.

Kouros, N. (2013). Australian women turning to home births after hospital trauma. *Monash Bioethics Review, 31*(2), 10-11

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*, 29. doi: 10.1186/s12884-015-0759-x

Martin, C. R., Hollins Martin, C. J., Burduli, E., Barbosa-Leiker, C., Donovan-Batson, C., & Fleming, S. E. (2016a). Measurement and structural invariance of the US version of the Birth Satisfaction Scale-Revised (BSS-R) in a large sample. *Women and Birth*. doi: 10.1016/j.wombi.2016.11.006

Martin, C. R., Vardavaki, Z., & Hollins Martin, C. J. (2016b). Measurement equivalence of the Birth Satisfaction Scale-Revised (BSS-R): further evidence of construct validity. *Journal of Reproductive and Infant Psychology*, 1-9. doi: 10.1080/02646838.2016.1184747

McDonald, S., Slade, P., Spiby, H., & Iles, J. (2011). Post-traumatic stress symptoms, parenting stress and mother-child relationships following childbirth and at 2 years postpartum. *Journal of Psychosomatic Obstetrics and Gynaecology, 32*(3), 141-146. doi: 10.3109/0167482X.2011.596962

International Consortium for Health Outcomes Measurement. (2016). *ICHOM Standard Set for Pregnancy and Childbirth*. http://www.ichom.org/medical-conditions/pregnancy-and-childbirth/.

Millsap, R. E. (1998). Group Differences in Regression Intercepts: Implications for Factorial Invariance. *Multivariate Behavioral Research, 33*(3), 403-424. doi: 10.1207/s15327906mbr3303_5

Quinn, E., Noble, J., Seale, H., & Ward, J. E. (2013). Investigating the potential for evidence-based midwifery-led services in very remote Australia: viewpoints from local stakeholders. *Women and Birth, 26*(4), 254

Ragupathy, R., Aaltonen, K., Tordoff, J., Norris, P., & Reith, D. (2012). A 3-dimensional view of access to licensed and subsidized medicines under single-payer systems in the US, the UK, Australia and New Zealand. *Pharmacoeconomics, 30*(11), 1051-1065. doi: 10.2165/11595270-000000000-00000

Razurel, C., & Kaiser, B. (2015). The role of satisfaction with social support on the psychological health of primiparous mothers in the perinatal period. *Women and Health, 55*(2), 167-186. doi: 10.1080/03630242.2014.979969

Redshaw, M., & Martin, C. R. (2009). Validation of a perceptions of care adjective checklist. *Journal of Evaluation in Clinical Practice, 15*(2), 281-288. doi: 10.1111/j.1365-2753.2008.00995.x

Renfrew, M. J., McFadden, A., Bastos, M. H., Campbell, J., Channon, A. A., Cheung, N. F., . . . Declercq, E. (2014). Midwifery and quality care: findings from a new evidence-informed framework for maternal and newborn care. *Lancet, 384*(9948), 1129-1145. doi: 10.1016/S0140-6736(14)60789-3

Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling. *Journal of Statistical Software, 48*(2), 1-36

Sandall, J., Soltani, H., Gates, S., Shennan, A., & Devane, D. (2016). Midwife-led continuity models versus other models of care for childbearing women. *Cochrane Database of Systematic Reviews, 4*, CD004667. doi: 10.1002/14651858.CD004667.pub5

Sawyer, A., Ayers, S., Abbott, J., Gyte, G., Rabe, H., & Duley, L. (2013). Measures of satisfaction with care during labour and birth: a comparative review. *BMC Pregnancy and Childbirth, 13*, 108. doi: 10.1186/1471-2393-13-108

Schmitt, N. (1996). Uses and abuses of coefficient alpha. *Psychological Assessment, 8*(4), 350-353

Sijtsma, K. (2009). On the use, the misuse, and the very limited usefulness of Cronbach's alpha. *Psychometrika, 74*(1), 107-120. doi: 10.1007/s11336-008-9101-0

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. doi: 10.1111/j.1744-6163.2009.00234.x

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

Team, R. C. (2017). *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing.

Tierney, O., Sweet, L., Houston, D., & Ebert, L. (2016). The Continuity of Care Experience in Australian midwifery education-What have we achieved? *Women and Birth.* doi: 10.1016/j.wombi.2016.10.006

Torchiano, M. (2015). *Effsize: Efficient Effect Size Computation.* https://CRAN.R-project.org/package=effsize.

Tracy, S. K., Hartz, D. L., Tracy, M. B., Allen, J., Forti, A., Hall, B., & , e. a. (2013). Caseload midwifery care versus standard maternity care for women of any risk: M@NGO, a randomised controlled trial. *Lancet, 382*(9906), 1723

Tracy, S. K., Welsh, A., Hall, B., Hartz, D., Lainchbury, A., Bisits, A., & al., e. (2014). Caseload midwifery compared to standard or private obstetric care for first time mothers in a public teaching hospital in Australia: a cross sectional study of cost and birth outcomes. *BMC Pregnancy Childbirth, 14*, 46

Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods, 3*(1), 4-70. doi: doi: 10.1177/109442810031002

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

Weaver, E. W., Clark, K. F., & Vernon, B. A. (2005). Obstetricians and midwives modus vivendi for current times. *Medical Journal of Australia, 182*(9), 436-437

Webster, J., Hall, L., Somville, T., Schneider, P., Turnbull, R., & Smith, P. (2006). Prospective testing of the Brisbane Postnatal Depression Index. *Birth, 33*(1), 56-63. doi: 10.1111/j.0730-7659.2006.00074.x

Yelland, J., Sutherland, G., & Brown, S. J. (2010). Postpartum anxiety, depression and social health: findings from a population-based survey of Australian women. *BMC Public Health, 10*, 771. doi: 10.1186/1471-2458-10-771

Zou, G. Y. (2007). Toward using confidence intervals to compare correlations. *Psychological Methods, 12*, 399–413. doi: doi: 10.1037/1082-989X.12.4.399

Table 1. Mean, standard deviation and distributional characteristics
of A-BSS-R sub-scales and total score. se = standard error of kurtosis.

Subscale	Mean	SD	Range	Skew	Kurtosis	se
Stress	10.63	3.45	16	-0.60	0.15	0.24
Attributes	5.43	1.91	8	-0.44	-0.44	0.14
Quality	14.34	2.08	9	-1.30	0.96	0.15
Total score	30.40	5.88	27	-0.39	-0.38	0.42

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

Item	Item content	Domain*	Mean	SD	Range	Skew	Kurtosis	se
BSS-R 1	I came through childbirth virtually unharmed	SE	3.08	1.04	4	-1.11	0.47	0.07
BSS-R 2	I thought my labour was excessively long	SE	2.71	1.21	4	-0.71	-0.53	0.09
BSS-R 3	The birthing room staff encouraged me to make decisions about how I wanted my birth to progress	QC	3.39	0.75	3	-0.84	-0.51	0.05
BSS-R 4	I felt very anxious during my labour and birth	WA	2.58	1.12	4	-0.45	-0.75	0.08
BSS-R 5	I felt well supported by staff during my labour and birth	QC	3.67	0.60	3	-1.75	2.49	0.04
BSS-R 6	The staff communicated well with me during labour	QC	3.62	0.63	3	-1.64	2.43	0.04
BSS-R 7	I found giving birth a distressing experience	SE	2.69	1.19	4	-0.64	-0.57	0.08
BSS-R 8	I felt out of control during my birth experience	WA	2.86	1.09	4	-0.76	-0.24	0.08
BSS-R 9	I was not distressed at all during labour	SE	2.15	1.17	4	0.07	-0.98	0.08
BSS-R 10	The delivery room was clean and hygienic	QC	3.67	0.61	3	-1.90	3.47	0.04

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

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

Scale combination	Australian r	UK r	Z	95% CI	p
Stress-Attributes	0.67	0.57	1.67	(-0.02 – 0.22)	0.10
Stress-Quality	0.23	0.26	0.33	(-0.21 – 0.15)	0.74
Attributes-Quality	0.32	0.35	0.35	(-0.20 – 0.14)	0.73
Total score-Stress	0.89	0.86	1.31	(-0.02 – 0.08)	0.19
Total score-Attributes	0.83	0.80	0.92	(-0.04 – 0.10)	0.36
Totals score-Quality	0.60	0.63	0.49	(-0.15 – 0.09)	0.62

Table 4. Comparison of BSS-R total and sub-scale scores differentiated by birth delivery type (after Vardavaki et al., 2015). Standard deviations are in parentheses, degrees of freedom = 196, CI = confidence interval. Non-normal delivery group comprised, Ceasarean section N=18, Emergency Ceasarean section N=28, Forceps N=9 and Vacuum N=14.

BSS-R Scale	Unassisted vaginal delivery (N=129)	Assisted/ Operative delivery (N=69)	95% CI	t	p	Hedges g	Hedges g 95% CI	Effect size
Stress	11.54 (3.07)	8.93 (3.48)	1.67 - 3.56	5.45	<0.001	0.81	0.50 - 1.11	Large
Attributes	5.72 (1.80)	4.90 (2.02)	0.27 - 1.37	2.94	0.004	0.44	0.14 - 0.73	Small
Quality	14.54 (1.92)	13.97 (2.32)	-0.04 - 1.17	1.83	0.07	0.27	-0.02 - 0.57	Small
Total score	31.80 (5.48)	27.80 (5.75)	2.36 - 5.64	4.81	<0.001	0.71	0.41 - 1.02	Medium

Table 5. Cronbach's alpha of BSS-R sub-scales and total score and comparison with original UK BSS-R (Hollins Martin and Martin, 2014). Degrees of freedom = 1.

Subscale	Australian alpha	UK alpha	χ^2	p
Stress	0.74	0.71	0.37	0.54
Attributes	0.66	0.64	0.06	0.81
Quality	0.81	0.74	3.07	0.08
Total score	0.81	0.79	0.43	0.51

Table 6. Invariance analysis of Australian (A-BSS-R) and UK (BSS-R) datasets.

Model	χ^2 (df)	Model comparison	$\Delta\chi^2$	Δdf	p	RMSEA	SRMR	CFI	ΔCFI	Different
1. Overall	42.66(32)	na	na	na	na	0.028	0.032	0.992	na	na
2. Configural	114.05(64)	na	na	Na	na	0.061	0.051	0.962	na	na
3. Metric	127.06(71)	2	13.01	7	0.07	0.061	0.058	0.958	0.004	No
4. Scalar	186.66(78)	3	59.60	7	<0.05	0.081	0.069	0.918	0.04	Yes
5. Partial scalar BSS-R 1. (intercepts)	152.01(77)	3	24.95	6	<0.05	0.068	0.062	0.943	0.015	Yes
6. Partial scalar BSS-R 1. & BSS-R 3.(intercepts)	141.29(76)	3	14.23	5	<0.05	0.064	0.061	0.951	0.007	No