Psychometric properties of a Saudi Arabian version of the Birth Satisfaction Scale-Revised (BSS-R)

Juliana Linnette D'Sa[1]* , Ibtesam Omar Jahlan[1] , Eman S. Alsatari[2] ,
Sahar Zamzam[1,3] , and Colin R. Martin[4]

[1] Maternal and Child Health Nursing Department, College of Nursing, King Saud University, Riyadh, Saudi Arabia
[2] School of Nursing, University of Louisville, Louisville, KY 40202, United States
[3] Faculty of Nursing, Alexandria University, Egypt
[4] Institute for Health and Wellbeing, University of Suffolk, Ipswich, Neptune Quay, IP4 1QJ, United Kingdom

*Corresponding author:
Juliana Linnette D'Sa. MN, MPhil, PhD Nursing, Professor, Maternal and Child Health Nursing Department, College of Nursing, King Saud University, Riyadh, Saudi Arabia
Email: jdsa@ksu.edu.sa

Abstract

Background: Maternal outcomes are closely associated with birth satisfaction, and the Birth Satisfaction Scale-Revised (BSS-R), a concise, multidimensional self-report measure, has undergone translation and validation internationally. However, research on birth satisfaction in Saudi Arabia is scarce. The absence of valid Arabic-language tools for the Saudi population may impede critical research on this topic, necessitating the translation and use of psychometrically sound instruments for measuring birth satisfaction in Saudi women.

Objective: This study aimed to translate and evaluate the psychometric properties of the Saudi Arabian version of the BSS-R (SA-BSS-R).

Methods: A total of 218 Saudi women participated in the study, and psychometric analysis of the translated SA-BSS-R involved confirmatory factor analysis, divergent validity analysis, and known-group discriminant validity assessment within a cross-sectional study design.

Results: The three-factor BSS-R measurement model displayed poor fit, and internal consistency fell below the threshold value. Additionally, it was observed that women undergoing an episiotomy had significantly lower overall SA-BSS-R scores.

Conclusion: The SA-BSS-R manifested atypical measurement properties in this population. Despite insightful observations related to episiotomy, the identified measurement shortcomings highlight the need for a more robust and culturally sensitive translation to enhance measurement characteristics.

Keywords

Background

Childbirth is a pivotal event in the lives of women. While some find it satisfying, the childbirth experience can be dissatisfying for others, significantly impacting their lives and relationships with the newborn. A negative birth experience is associated with postpartum depression, posttraumatic stress disorder (Ayers et al., 2016; Bell & Andersson, 2016), fear of childbirth (Størksen et al., 2013), an increased inclination towards cesarean delivery (Fenwick et al., 2010), delayed subsequent pregnancy (Shorey et al., 2018), and a reluctance towards future childbirth (Nahaee et al., 2020).

Conversely, a positive childbirth experience can be empowering and satisfying (Hosseini Tabaghdehi et al., 2020), correlating with positive outcomes like an increased tendency to breastfeed (Davis & Sclafani, 2022), improved mother-child bonding, maternal caregiving attitude and behavior (Bell et al., 2018), and subsequently, high levels of satisfaction (Fair & Morrison, 2012). Quality intrapartum, respectful, and humane care empowers women and fosters satisfaction with their birth

experiences (Jahlan et al., 2016). Impressions of birth experiences endure over time and can significantly impact the lives of women (Bossano et al., 2017).

Birth satisfaction, reflecting a woman's overall appraisal of her birthing experience, is a complex construct consisting of three elements: stress experienced (SE), personal characteristics (PC), and the quality of care (QC) received (Hollins Martin & Martin, 2014). Assessing women's evaluations of their birth experiences and the factors influencing satisfaction levels is implicitly critical. Low birth satisfaction often stems from birth characteristics, such as obstetrical interventions and the mode of delivery, common in medicalized models. In such models, women may lose autonomy over birth decisions (Scamell et al., 2017) and frequently undergo interventions and procedures that may disrupt the natural childbirth process (World Health Organization, 2018). Despite international recommendations favoring episiotomy only when clinically indicated, episiotomy rates remain high in the Gulf countries (52%), including Saudi Arabia (45%), with subjective indications (Al-Zabidi et al.,

2021). Episiotomy is associated with low birth satisfaction and has negative physical and mental health impacts on women (Mohammad et al., 2014a). Postpartum perineal pain, discomfort, and anxiety associated with episiotomy can hinder women's enjoyment of childbirth and affect their mental health (He et al., 2020). The fear and negative experiences related to episiotomy can impact long-term confidence in sexuality and their decisions regarding future deliveries (He et al., 2020).

However, episiotomy is likely to engender negative sentiments about childbirth, leading to dissatisfaction (Calik et al., 2018; Mohammad et al., 2014b; Nahaee et al., 2020). Moreover, women undergoing instrumental or emergency cesarean deliveries tend to report lower levels of satisfaction with their birthing experiences (Viirman et al., 2022). In contrast, those who experience spontaneous vaginal deliveries tend to express a more positive assessment of their birthing experiences (Hildingsson et al., 2013). Measuring satisfaction levels holds fundamental importance given the various factors influencing the childbirth experience.

For a thorough and reliable evaluation of women's assessments of their childbirth experiences, it is imperative to employ self-report measures with sound psychometric properties. However, previous reviews have consistently found that most instruments measuring birth satisfaction lack robust psychometric properties and theoretical coherence (Al Nadabi & Mohammed, 2019; Sawyer et al., 2013). A review of 18 years (2000–2018) identified nine instruments used in Arabic-speaking nations, none of which were applied in Saudi Arabia. The conclusion was that most of these instruments lacked robust psychometric properties (Al Nadabi & Mohammed, 2019). Theoretical coherence, validity, and reliability are essential characteristics of measurement instruments (Sawyer et al., 2013).

The Birth Satisfaction Scale-Revised (BSS-R), a psychometrically sound, theoretically grounded, and widely used three-dimensional instrument (Hollins Martin & Martin, 2014), is considered the most appropriate tool for evaluating women's satisfaction with childbirth. Its brevity is a notable advantage, making it a preferred instrument. The BSS-R gained significant traction after the International Consortium for Health Outcomes Measurement recommended its global use in 2017 (The International Consortium for Health Outcome Measurement, 2017). Initially developed in the English language in the United Kingdom, the BSS-R has been translated into numerous languages and extensively utilized in 68 countries and over 270 sites worldwide, with these figures continually increasing (Hollins Martin & Martin, 2024). The most recent addition is the Indian version of the BSS-R (Tiwari et al., 2023). While a few Middle Eastern countries have translated and validated the instrument, including a Persian version (Mortazavi et al., 2021), a psychometrically tested and culturally adapted Saudi Arabian version is currently unavailable. Although an unpublished master's thesis reported the translation and use of the BSS-R among Saudi women (Almalki, 2021), the instrument lacked robust translation and validation procedures, potentially limiting its utility in future studies. Given the scarcity of research on birth satisfaction in Saudi Arabia and the absence of valid and reliable Arabic-language instruments for this population, there is an urgent need for psychometrically assessed instruments to measure birth satisfaction in Saudi women.

Acknowledging the imperative to enhance healthcare quality, Saudi Arabia introduced the National Transformation Policy in alignment with Vision 2030 (Saudi Vision, 2023). The Health Sector Transformation Program is founded on four key elements: better care, sustainability, and workforce. Maternity and childcare represent two of the six pathways designed to enhance health. The New Model of Care program also seeks to facilitate safe births for women (Saudi Vision, 2023). While Saudi Arabia encourages institutional births following a predominantly medicalized model, there is a vision to shift maternity care from a medicalized approach to midwifery-led care for uncomplicated home births (Ministry of Health, 2021). This underlines the importance of using robust quality measures to evaluate women's satisfaction with their birthing experiences.

The Ministry of Health in Saudi Arabia primarily emphasizes providing quality care while transforming the healthcare sector. Women's satisfaction with their birth experiences is deemed a crucial measurable outcome in evaluating the quality of care during labor and birth (Nilver et al., 2017). Despite this, birth satisfaction has not been extensively explored in Saudi Arabia. The absence of valid and reliable Arabic-language instruments may have hindered research on this subject. There is a pressing need for an Arabic-language, culturally adapted, and psychometrically tested instrument to assess birth satisfaction as a metric for evaluating the quality of care among Saudi women.

The objectives were to evaluate the three-dimensional measurement model of the BSS-R within the translated Saudi Arabian version (SA-BSS-R), to evaluate the internal consistency of the SA-BSS-R: quality of care (QC), women's attribute (WA), and stress experienced (SE) during childbirth subscales and the overall SA-BSS-R scale, to assess the known group discriminant validity of the SA-BSS-R and its divergent validity, and to examine the differences between the SA-BSS-R score in relation to episiotomy.

Methods

Study Design

The study employed a cross-sectional design. The BSS-R consists of ten items, evaluating three domains of the birth experience: SE with four items, WA with two items, and QC with four items, all utilizing a self-report format (Hollins Martin & Martin, 2014). Higher subscales and total scale scores indicate greater satisfaction with the specific domains or the overall birth experience. During Phase 1, the translation and cross-cultural adaptation of the BSS-R into Arabic followed the recommendations of Beaton et al. (2000). In Phase 2, a psychometric evaluation was carried out.

Phase 1: Translation of the BSS-R

Stage 1: Translation. The translation process was spearheaded by the second author, who possessed a doctorate in midwifery and was a native Arabic speaker. Bilingual language translators were employed to convert the English version into Arabic, comprising a health professional with substantial expertise in the concept of birth satisfaction and a nonprofessional proficient in both English and Arabic.

Stage 2: Synthesis. The two translated versions were synthesized into a single tentative Arabic version. Two items, 'I found giving birth a distressing experience' (Item 7) and 'I

was not distressed at all during labor' (item 9), differed in the translation, which was synthesized into a common and culturally acceptable statement after the translator consensus.

Stage 3: Back-translation. The tentative Arabic version was back-translated independently by two translators whose native language was English, who were proficient in Arabic and were unaware of the concept being explored. Both back-translated versions were synthesized to create a single version. Furthermore, discrepancies between the two back-translators in the choice of Arabic words for four items were resolved, and the most appropriate words were chosen based on translator consensus.

Stage 4: Expert committee review. The original BSS-R, tentative Arabic version, and synthesized back-translated version were presented to five experts to assess content validity and cultural relevance. Two practicing midwives, two faculty members (one with a specialty in midwifery and the other in psychiatric nursing), and a linguist with postgraduate experience constituted the expert committee. They assessed the content relevance of the 10 Arabic items on a four-point scale (1 = not relevant, 2 = somewhat relevant, 3 = quite relevant, and 4 = very relevant), as well as instructions and response sets for appropriate wording. Cross-cultural validation was performed by assessing items on a binary scale (yes/no) for semantic, idiomatic, experiential, and conceptual equivalence. The most culturally appropriate words were used for four items: 'I felt supported by staff during my labor and birth' (item 5), 'the staff communicates well during labor' (item 6), 'I found giving birth a distressing experience' (item 7) and 'I was not distressed at all during labor' (item 9). For example, the Arabic word that translated to 'suffer' was deemed more appropriate rather than the word for 'distress' and was accordingly used in the translation process. After discussion and critical examination, the panel agreed on the pre-final Arabic version of the scale. The item- and scale-level content validity indices were computed after recoding the four-point scale to create a binary outcome: 1 and 2 (low relevance) and 3 and 4 (high relevance). The item content validity index (I-CVI) of the ten items ranged from 0.6 (for items 3 and 7) to 1.0, and the scale content validity index (S-CVI) was 0.92. For Item 3, two experts suggested replacing the Arabic word for 'labor' with 'delivery'; for Item 7, the Arabic word for 'painful' instead of the one for 'distressing' was suggested to be used. Hence, the CVI was 0.62 for the two items. The modifications were made accordingly.

Stage 5: Pretesting. The pre-final Arabic version was pre-tested on 30 postnatal women who qualitatively evaluated the instrument for readability, feasibility, and acceptance. No modifications were made to the instrument.

Phase 2: Psychometric Evaluation

Data were gathered between November 2021 and February 2022 by the third author and trained assistants from willing participants. The study's objective was elucidated, and written consent was secured. The BSS-R was administered to 247 women selected conveniently from the postnatal ward within seven days of giving birth. The completed scales were returned directly to the researcher or the trained assistant. The final dataset comprised 218 participants.

Data Analysis

Statistical analysis was undertaken using R (R Core Team, 2022), Lavaan (Rosseel, 2012), SemTools (Jorgensen et al., 2022), Cocron (Diedenhofen & Musch, 2016) and Cocor (Diedenhofen & Musch, 2015) packages.

Construct validity. Confirmatory factor analysis (CFA) was utilized to assess the established three-dimensional measurement model of the BSS-R. Prior research on the translation and validation of the BSS-R has also employed CFA (Barbosa-Leiker et al., 2015; Burduli et al., 2017; Emmens et al., 2023; Hollins Martin & Martin, 2014; Hollins Martin & Martin, 2024; Nakic Rados et al., 2023; Ratislavová et al., 2024). Initially, data underwent screening to ensure that distributional characteristics met the parametric requirements for CFA (Brown, 2015), including evaluating individual item skew and kurtosis and excluding multivariate outliers (Kline, 2000). The BSS-R comprises three correlated factors represented by the Stress Experienced (SE), Women's Attributes (WA), and Quality of Care (QC) subscales as a measurement model (Hollins Martin & Martin, 2014). Additionally, a bifactor model within the BSS-R was explored, differentiating a general factor of birth experience from specific factors and determining the degree of variance explained by general and specific factors (Martin et al., 2018). The maximum-likelihood approach was employed for model estimation (Brown, 2015; Kline, 2011), and the fit to data was evaluated using the comparative fit index (CFI) (Bentler, 1990), root mean squared error of approximation (RMSEA) (Steiger & Lind, 1980), and square root mean residual (SRMR) (Hu & Bentler, 1999). Conventional threshold values of >0.90 (CFI), <0.08 (RMSEA), and <0.06 (SRMR) were adopted to determine the satisfactory fit of the model to the data.

Internal consistency. The assessment of internal consistency for utilized Cronbach's alpha (Cronbach, 1951). A threshold of 0.70 or higher was applied to signify acceptable internal consistency. For the two-item Women's Attributes (WA) subscale, internal consistency was determined through inter-item correlation, with the established threshold range convention of 0.15–0.50 (Clark & Watson, 1995).

Known group discriminant validity. In previous translation and validation studies of the BSS-R, delivery type has been commonly used to assess known group discriminant validity, typically associating unassisted delivery with higher birth satisfaction (Hollins Martin & Martin, 2014; Romero-Gonzalez et al., 2019). Recent studies have examined specific delivery aspects such as unassisted vaginal delivery, assisted vaginal delivery, emergency cesarean section, and elective cesarean section (Emmens et al., 2023; Nakic Rados et al., 2023; Ratislavová et al., 2024). These studies (with some variability) showed minimal differences between unassisted vaginal delivery and elective cesarean section but notably lower BSS-R scores in cases involving assisted vaginal delivery (forceps or ventouse) or emergency cesarean section. Given the low incidence of assisted vaginal delivery and emergency cesarean section in the study population, the current study tested the hypothesis of no significant difference between groups dichotomized based on either unassisted vaginal delivery or elective cesarean section. To assess known group discriminant validity by comparing groups where birth experiences were anticipated to differ, women who underwent episiotomy were compared with those who did not. We

hypothesized that women who underwent episiotomy would have significantly lower scores on the SE subscale and overall BSS-R. Additionally, we predicted no score differences between the groups on the WA and QC subscales.

Divergent validity. It was assessed using correlation coefficients (Pearson's *r*) between participant age and the overall SA-BSS-R score and individual subscale scores.

Ethical Considerations
This study was approved by the Institutional Review Board of the King Saud University, College of Medicine (No. 21/0868/IRB). The Nursing Affairs of King Saud University Medical City provided permission for data collection. Informed written consent was obtained from participants; participation was voluntary. Data confidentiality was maintained. Permission to translate the BSS-R was taken from the author of the instrument.

Results

Demographic Characteristics of the Participants
Of the 242 participants, 18 non-Saudi participants were excluded from the dataset. Multivariate outliers (*n* = 6) were removed by calculating the Mahalanobis distances; thus, a final dataset comprising 218 participants was available for psychometric appraisal. The average age of the participants was 30.11 (SD = 6.07). Of the participants, 61 (28%) were primiparas. The mean number of days since delivery was 1.67 (SD = 0.99; range = 0–6 days). The average length of gestation was 38.01 (SD = 1.80) weeks. Episiotomy was performed on 72 women (33%), and 81 women (37%) attended childbirth classes. A majority of the participants (*N* = 129, 59%) had planned pregnancies. One hundred and thirty four (61%) women had an unassisted vaginal delivery, while 9 (4%), 11 (5%), and 3 (1%) women had forceps, ventouse, or breech delivery, respectively. Sixty women (28%) underwent a planned cesarean section, while one woman (<1%) underwent an emergency cesarean section.

Descriptive and Distributional Characteristics of the SA-BSS-R
No excessive skew or kurtosis was observed in individual SA-BSS-R items, subscales, or total scores (**Table 1**).

Table 1 Descriptive and distributional characteristics of the SA-BSS-R

Item	Item content	Domain	Mean	SD	Min-Max	Skew	Kurtosis	Standard Error
BSS-R 1	I came through childbirth virtually unscathed	SE	2.82	1.09	0-4	-0.91	0.07	0.07
BSS-R 2	I thought my labor was excessively long	SE	1.70	1.23	0-4	0.07	-1.15	0.08
BSS-R 3	The delivery room staff encouraged me to make decisions about how I wanted my birth to progress	QC	3.07	1.12	0-4	-1.07	0.14	0.08
BSS-R 4	I felt very anxious during my labor and birth	WA	0.98	1.05	0-4	1.02	0.22	0.07
BSS-R 5	I felt well supported by staff during my labor and birth	QC	3.39	0.85	0-4	-1.55	2.28	0.06
BSS-R 6	The staff communicated well with me during labor	QC	3.40	0.84	0-4	-1.65	2.73	0.06
BSS-R 7	I found giving birth a distressing experience	SE	0.84	1.02	0-4	1.08	0.30	0.07
BSS-R 8	I felt out of control during my birth experience	WA	1.89	1.27	0-4	-0.10	-1.24	0.09
BSS-R 9	I was not distressed at all during the labor	SE	1.14	1.12	0-4	0.87	-0.05	0.08
BSS-R 10	The delivery room was clean and hygienic	QC	3.56	0.76	0-4	-2.05	4.46	0.05
Stress	Sub-scale total		6.50	2.67	0-15	0.69	0.49	0.18
Attributes	Sub-scale total		2.87	1.95	0-8	0.28	-0.71	0.13
Quality	Sub-scale total		13.42	3.02	0-16	-1.34	1.76	0.20
Total	Total score		22.79	5.07	5-37	0.17	0.79	0.34

Note: *Domain of the Saudi Arabian BSS-R. SE = Stress experienced during childbearing, WA = Women's attributes, QC = Quality of Care
Legend: Mean, SD standard deviation, and distributional characteristics of individual Saudi Arabian BSS-R items, sub-scale totals, and the total Saudi Arabian BSS-R score

Confirmatory Factor Analysis
The evaluated factor models are outlined in **Table 2**. The single-factor model (Model 1) had a poor fit to the data, as anticipated. The conventional three-factor measurement model (Model 2) in the BSS-R also demonstrated a poor fit to the data. Similarly, the bifactor model (Model 3) indicated a poor fit for the data. Investigation of modification indices for Model 2 suggested that the model fit could be enhanced by specifying Item 1 to load on the QC factors. This model (Model 4) resulted in an acceptable fit to the data in terms of the RMSEA and CFI indices but a mediocre fit to the SRMR index (see **Figure 1**). Finally, a nine-item, three-factor model was tested, excluding Item 1 (Model 5). Once again, the model fit was acceptable according to RMSEA and CFI but mediocre regarding the SRMR.

Table 2 Confirmatory factor analysis of the Saudi Arabian BSS-R

Model	χ^2 (df)	χ^2/df	*p*	RMSEA	SRMR	CFI
1. Single factor	257.37 (35)	7.35	<0.001	0.171	0.159	0.67
2. Three-factor	111.55 (32)	3.49	<0.001	0.107	0.112	0.884
3. Bifactor	93.54 (26)	3.60	<0.001	0.109	0.106	0.901
4. Modified three-factor	73.70 (32)	2.30	<0.001	0.077	0.070	0.939
5. Three-factor minus item 1	66.15 (24)	2.76	<0.001	0.090	0.072	0.935

Note: In Model 3. WA items were set to be equal in relation to contemporary practice for the run of bifactor models.
Without this constraint, the model fit of the bifactor model was similar: χ^2 = 93.54, df = 26, RMSEA = 0.109, SRMR = 0.106, CFI = 0.901
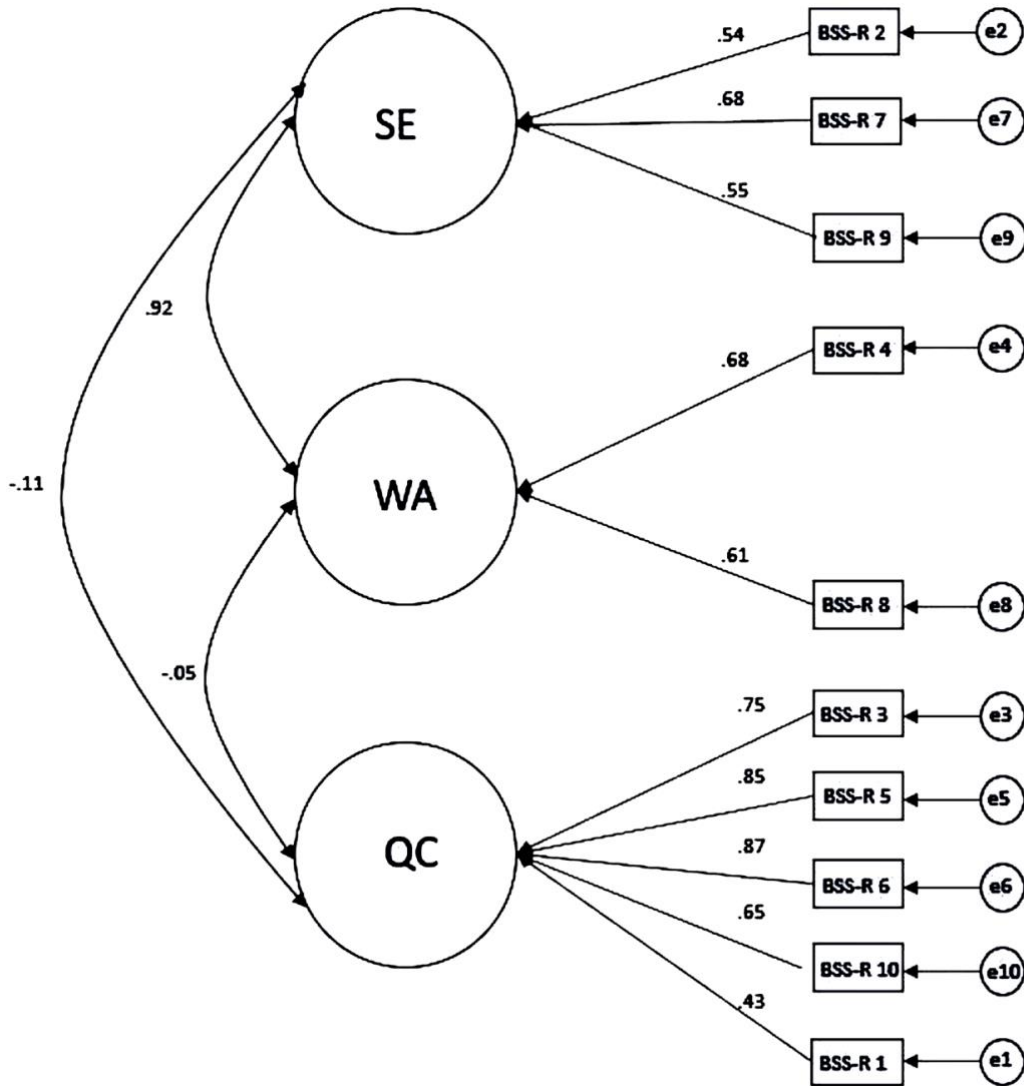
Figure 1 Modified three-factor model (Model 4) with standardized item-factor loadings and correlations between factors indicated. Only the correlation between SE and WA factors is substantive within the model

SA-BSS-R Subscale and Total Score Correlations

By employing a method outlined by Diedenhofen and Musch (2015) and excluding non-significant correlation comparisons between the SE and WA subscales, the QC subscale, and the overall SA-BSS-R score, it was observed that all correlation combinations between the original UK study (Hollins Martin & Martin, 2014) and the present study were significantly different. Furthermore, the correlations between the scales were found to be lower than those reported in the original UK study (p <0.05) (**Table 3**).

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

Scale Combination	Current Study r	UK Study r	Z	95% CI	p
Stress-Attributes	0.50	0.57	1.03	(-0.06 – 0.20)	0.30
Stress-Quality	0.07	0.26	2.06	(0.01 – 0.37)	0.04
Attributes-Quality	0.05	0.35	3.31	(0.12 – 0.47)	<0.001
Total score-Stress	0.76	0.86	3.12	(0.04 – 0.17)	0.002
Total score-Attributes	0.62	0.80	3.92	(0.09 – 0.28)	<0.001
Totals score-Quality	0.61	0.63	0.34	(-0.10 – 0.14)	0.73

Internal Consistency

The Cronbach's alpha for the SE subscale of the SA-BSS-R fell below the acceptable criterion (0.40). Upon recalculating Cronbach's alpha and excluding Item 1, which was identified as problematic in the CFA, there was an improvement in the alpha value, although it remained suboptimal at 0.60. The alpha value for the overall scale was also suboptimal at 0.64, while Cronbach's alpha for the QC subscale was excellent at 0.86. The correlation between the two WA subscale items was acceptable, with a r value of 0.41. In a final attempt, the alpha value was recalculated for the scale after excluding Item 1, but it remained suboptimal at 0.64.

Known-Group Discriminant Validity

No statistically significant differences were noted in the SE and WA subscales or the overall score among different delivery types. However, a statistically significant difference was observed in the QC subscales. Individuals who had an unassisted vaginal delivery had significantly higher scores compared to those who underwent elective cesarean section, although with a small effect size (**Table 4**).

Table 4 Comparison of Saudi Arabian BSS-R total and sub-scale scores differentiated by delivery type

BSS-R Scale	Vaginal Delivery (*N* = 134)	Elective Caesarean Section (*N* = 60)	*t*	*p*	Hedge's g	(95% CI)	Effect Size
Stress	6.29 (2.66)	6.88 (2.74)	1.42	0.16	0.22	-0.09 – 0.53	Small
Attributes	2.72 (1.91)	3.00 (2.02)	0.91	0.36	0.14	-0.16 – 0.45	Negligible
Quality	13.90 (2.65)	13.07 (3.02)	1.93	0.05	0.30	-0.01 – 0.61	Small
Total score	22.92 (4.88)	22.95 (5.75)	0.04	0.97	<0.01	-0.30 – 0.31	Negligible

Note: Standard deviations are in parentheses; degrees of freedom = 192

When comparing women categorized by episiotomy status, significant differences emerged between the groups in the SE subscale and the overall SA-BSS-R score, indicating that women who underwent episiotomy had lower scores (**Table 5**). However, no statistically significant differences were noted between the groups on WA or QC subscales.

Table 5 Comparison of Saudi Arabian BSS-R total and sub-scale scores differentiated by episiotomy status

BSS-R Scale	Episiotomy (*N* = 72)	No Episiotomy (*N* = 137)	*t*	*p*	Hedge's g	(95% CI)	Effect Size
Stress	6.04 (2.38)	6.82 (2.76)	2.04	0.04	0.30	0.01 – 0.58	Small
Attributes	2.57 (1.85)	3.03 (1.97)	1.64	0.10	0.24	-0.05 – 0.52	Small
Quality	13.17 (3.52)	13.61 (2.71)	1.00	0.32	0.15	-0.14 – 0.43	Negligible
Total score	21.78 (4.70)	23.46(5.18)	2.30	0.02	0.33	0.04 – 0.62	Small

Note: Standard deviations are in parentheses, degrees of freedom = 207

Divergent Validity

There were no significant correlations observed between participants' age and SE ($r = 0.08$, p = 0.21), WA ($r = 0.07$, *p* = 0.31), and QC ($r = 0.004$, *p* = 0.96) subscales, as well as the overall total score of SA-BSS-R ($r = 0.07$, *p* = 0.28).

Discussion

The findings of our study are somewhat disappointing regarding overall validation and equivalence, particularly concerning the comparability between the SA-BSS-R and the original UK version (Hollins Martin & Martin, 2014). Numerous translation studies have replicated the three-dimensional model of the BSS-R in country-specific versions (Barbosa-Leiker et al., 2015; Burduli et al., 2017; Emmens et al., 2023; Hollins Martin & Martin, 2014; Hollins Martin & Martin, 2024; Nakic Rados et al., 2023; Ratislavová et al., 2024; Romero-Gonzalez et al., 2019; Skvirsky et al., 2020). We observed studies that reported a poor fit to the three-dimensional measurement model but an excellent fit to alternative models; for instance, the Indian version by Tiwari et al. (2023) demonstrated an excellent fit to a two-factor model, excluding the two WA items.

However, within the context of our investigation, we could not identify a compelling fit to the data for the three-dimensional measurement model, bifactor model, or modified alternative versions. Therefore, it appears necessary to conclude that an additional revision of the SA-BSS-R may be required to establish a satisfactory fit for the data of the three-dimensional measurement model. Alternatively, the three-dimensional measurement model of the BSS-R might not align conceptually with the birthing culture in Saudi Arabia. This is surprising considering that the Persian version, validated for use in Iran, offered a good fit for the data for the three-dimensional model (Mortazavi et al., 2021).

Balancing the possibilities, we suspect a key issue could be the translation of Item 1. The expert panel noted that translating this item was particularly challenging using the forward and backward translation processes, and there was disagreement among the experts initially regarding the accuracy of the translation of this specific item. Additionally, removing this item significantly improved Cronbach's alpha for the SE subscale, suggesting limited conceptual commonality between Item 1 and the other three SE subscale items.

Although the modification indices indicated that Item 1 might be more appropriately specified as loading on the quality factor from a statistical perspective, a compelling theoretical justification for suggesting such a realignment is elusive and cannot be justified by any rationale within the existing literature. The research team must also acknowledge additional potential limitations in the integrity of this translation, which are not limited exclusively to Item 1. The Cronbach's alpha for the entire scale was below the acceptable threshold. We also observed that the correlations between the subscales were generally lower than those in the original BSS-R. Considering these findings, we acknowledge that a full review of all the BSS-R SE subscale items and a further round of translation and back translation is likely required to determine the fit to data of the three-dimensional measurement model or evaluate the possibility that the three-dimensional measurement model is culturally bound to a Western context, although this seems less likely given the support for the three-dimensional measurement model in non-Western contexts (Mortazavi et al., 2021).

By contrast, it was observed that the alpha of the QC subscale was very good; thus, further revision of the four items that make up this particular subscale may be unnecessary.

Additionally, no problems were observed with the two WA items regarding the correlation between items, which were within the limits of Clark and Watson (1995).

The known-group discriminant analysis provided additional evidence indicating minimal differences in the birth experience between women who had an unassisted vaginal delivery and those who opted for elective cesarean section. However, women who experienced unassisted vaginal delivery reported a significantly more positive birth experience on the QC subscale of the SA-BSS-R compared to those who underwent elective cesarean section. In contrast, Ratislavová et al. (2024) found no distinctions between unassisted vaginal delivery and elective cesarean section across any of the BSS-R subscales or the overall score. This suggests that, within the Saudi context, women may perceive better care when they have an unassisted vaginal delivery. Nevertheless, considering the translation limitations, it is possible that these findings might be influenced by the translation process, particularly within the QC subscale, although this seems unlikely given the satisfactory psychometric performance of the QC subscale.

Finally, in line with recent studies (Calik et al., 2018; Nahaee et al., 2020), our findings indicate that women who underwent an episiotomy expressed lower satisfaction with their birth experience than those who did not. Noteworthy is the significant divergence between these groups on the SE subscale of the SA-BSS-R, suggesting that women who had an episiotomy perceived childbirth as relatively more stressful. The topic of episiotomy remains insufficiently explored regarding its association with and impact on the birth experience; hence, further research is warranted.

Conclusion

Despite the identification of potentially significant distinctions between groups based on episiotomy status, the limitations in the current SA-BSS-R measurements imply the necessity for additional revisions. Consequently, the research team is undertaking a follow-up study to translate the BSS-R again, with a particular focus on items that may be affected by cultural factors within this population, to develop a psychometrically acceptable version of the measure.

Declaration of Conflicting Interest
There are no conflicts of interest to declare.

Funding
None.

Acknowledgment

Authors' Contributions
All authors contributed to the study's conception and design and approved the manuscript's final version. JLD and IOJ were involved in translating the instrument, ESA collected data, SZ performed data entry, and CRM analyzed the data and interpreted findings. JLD and CRM drafted the manuscript and reviewed the final version critically.

Authors' Biographies
Juliana Linnette D'Sa, MN, MPhil, PhD N, is a Professor in the Maternal and Child Health Nursing Department at the College of Nursing, King Saud University, Saudi Arabia. She has published original articles in the area of Maternal Nursing.

Ibtesam Omar Jahlan, PhD, Midwifery, is an Assistant Professor and Assistant Head of the Maternal and Child Health Nursing Department at the College of Nursing at King Saud University, Saudi Arabia. She is a consultant midwife and the Vice President of the Saudi Midwifery Association.

Eman S. Alsatari, MSN, PhD N Student, was a Nurse Educator at King Saud University Medical City, Saudi Arabia, at the time of the study. She is currently a Graduate Research Assistant, pursuing a Doctor of Nursing Philosophy degree at the School of Nursing, University of Louisville, USA. She has co-authored articles on neonatal health, frailty, and infection control and is a peer reviewer of several journals.

Sahar Zamzam, PhD N, is an Assistant Professor at King Saud University, Saudi Arabia, Faculty of Nursing, and Alexandria University, Egypt. She has published articles in the area of Maternal Nursing and Midwifery.

Prof. Colin. R. Martin, RN, BSc, MSc, MBA, PhD, YCAP, FHEA, CPsychol, CSci, AFBPsS, is the Clinical Director of the Institute of Health and Wellbeing at the University of Suffolk, UK, and Professor of Clinical Psychobiology and Applied Psychoneuroimmunology. He is a Chartered Health Psychologist and a Chartered Scientist. He has published numerous original research papers in the area of perinatal wellbeing.

Data Availability

Declaration of Use of AI in Scientific Writing
There is nothing to declare.

References

Al Nadabi, W., & Mohammed, M. A. (2019). Arabic language surveys measuring mothers' satisfaction during childbirth: A review. *Global Journal of Health Science*, *11*(6), 169. https://doi.org/10.5539/gjhs.v11n6p169

Al-Zabidi, F. Z. M., Elsirafi, M. S., Muqresh, M. A., Ali, M. A., & Billah, S. M. B. (2021). Situation analysis of episiotomy in the gulf: A scoping review. *International Journal of Human and Health Sciences*, *5*(3), 286-291.

Almalki, H. H. (2021). *Assessment of women's birth satisfaction* [Unpublished master's thesis, King Saud University]. Riyadh, Saudi Arabia.

Ayers, S., Bond, R., Bertullies, S., & Wijma, K. (2016). The aetiology of post-traumatic stress following childbirth: A meta-analysis and theoretical framework. *Psychological Medicine*, *46*(6), 1121-1134. https://doi.org/10.1017/S0033291715002706

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

Beaton, D. E., Bombardier, C., Guillemin, F., & Ferraz, M. B. (2000). Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine*, *25*(24), 3186-3191. https://doi.org/10.1097/00007632-200012150-00014

Bell, A. F., & Andersson, E. (2016). The birth experience and women's postnatal depression: A systematic review. *Midwifery*, *39*, 112-123. https://doi.org/10.1016/j.midw.2016.04.014

Bell, A. F., Andersson, E., Goding, K., & Vonderheid, S. C. (2018). The birth experience and maternal caregiving attitudes and behavior: A systematic review. *Sexual & Reproductive Healthcare*, *16*, 67-77. https://doi.org/10.1016/j.srhc.2018.02.007

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

Bossano, C. M., Townsend, K. M., Walton, A. C., Blomquist, J. L., & Handa, V. L. (2017). The maternal childbirth experience more than a decade after delivery. *American Journal of Obstetrics and Gynecology*, *217*(3), 342.e341-342.e348. https://doi.org/10.1016/j.ajog.2017.04.027

Brown, T. (2015). *Confirmatory factor analysis for applied research* (2nd ed.). New York: Guilford Press.

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

Calik, K. Y., Karabulutlu, Ö., & Yavuz, C. (2018). First do no harm-interventions during labor and maternal satisfaction: A descriptive cross-sectional study. *BMC Pregnancy and Childbirth*, *18*, 415. https://doi.org/10.1186/s12884-018-2054-0

Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment*, *7*(3), 309–319. https://doi.org/10.1037/1040-3590.7.3.309

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

Davis, A. M. B., & Sclafani, V. (2022). Birth experiences, breastfeeding, and the mother-child relationship: Evidence from a large sample of mothers. *Canadian Journal of Nursing Research*, *54*(4), 0844562122 21089475. https://doi.org/10.1177/08445621221089475

Diedenhofen, B., & Musch, J. (2015). Cocor: A comprehensive solution for the statistical comparison of correlations. *PloS One*, *10*(4), e0121945. https://doi.org/10.1371/journal.pone.0121945

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.

Emmens, B., Hollins Martin, C. J., & Martin, C. R. (2023). Translation and validation of the Dutch version of the Birth Satisfaction Scale-Revised (BSS-R). *Journal of Reproductive and Infant Psychology*, *41*(2), 213-227. https://doi.org/10.1080/02646838.2021.1979200

Fair, C. D., & Morrison, T. E. (2012). The relationship between prenatal control, expectations, experienced control, and birth satisfaction among primiparous women. *Midwifery*, *28*(1), 39-44. https://doi.org/10.1016/j.midw.2010.10.013

Fenwick, J., Staff, L., Gamble, J., Creedy, D. K., & Bayes, S. (2010). Why do women request caesarean section in a normal, healthy first pregnancy? *Midwifery*, *26*(4), 394-400. https://doi.org/10.1016/j.midw.2008.10.011

He, S., Jiang, H., Qian, X., & Garner, P. (2020). Women's experience of episiotomy: A qualitative study from China. *BMJ Open*, *10*(7), e033354. https://doi.org/10.1136%2Fbmjopen-2019-033354

Hildingsson, I., Johansson, M., Karlström, A., & Fenwick, J. (2013). Factors associated with a positive birth experience: An exploration of Swedish women's experiences. *International Journal of Childbirth*, *3*(3), 153-164. https://doi.org/10.1891/2156-5287.3.3.153

Hollins Martin, C. J., & Martin, C. R. (2014). Development and psychometric properties of the Birth Satisfaction Scale-Revised (BSS-R). *Midwifery*, *30*(6), 610-619. https://doi.org/10.1016/j.midw.2013.10.006

Hollins Martin, C. J., & Martin, C. R. (2024). *The Birth Satisfaction Scale-Revised (BSS-R)*. www.bss-r.co.uk

Hosseini Tabaghdehi, M., Keramat, A., Kolahdozan, S., Shahhosseini, Z., Moosazadeh, M., & Motaghi, Z. (2020). Positive childbirth experience: A qualitative study. *Nursing Open*, *7*(4), 1233-1238. https://doi.org/10.1002/nop2.499

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

Jahlan, I., Plummer, V., McIntyre, M., & Moawed, S. (2016). What women have to say about giving birth in Saudi Arabia. *Middle East Journal of Nursing*, *10*(1), 1-9.

Jorgensen, T. D., Pornprasertmanit, S., Schoemann, A. M., & Rosseel, Y. (2022). *semTools: Useful tools for structural equation modeling*. https://CRAN.R-project.org/package=semTools

Kline, P. (2000). *A psychometric primer*. London: Free Association Books.

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

Martin, C. R., Hollins Martin, C. J., Burduli, E., Barbosa-Leiker, C., Donovan-Batson, C., & Fleming, S. E. (2018). The Birth Satisfaction Scale–Revised (BSS-R): should the subscale scores or the total score be used? *Journal of Reproductive and Infant Psychology*, *36*(5), 530-535. https://doi.org/10.1080/02646838.2018.1490498

Ministry of Health. (2021). *Saudi home birth standards* https://www.moh.gov.sa/en/Ministry/MediaCenter/Publications/Pages/The-Saudi-Birth-Centre-Standards.pdf

Mohammad, K. I., Alafi, K. K., Mohammad, A. I., Gamble, J., & Creedy, D. (2014a). Jordanian women's dissatisfaction with childbirth care. *International Nursing Review*, *61*(2), 278-284. https://doi.org/10.1111/inr.12102

Mohammad, K. I., Shaban, I. A., Homer, C. S., & Creedy, D. (2014b). Women's satisfaction with hospital-based intrapartum care: A Jordanian study. *International Journal of Nursing and Midwifery*, *6*(3), 32-39. https://doi.org/10.5897/IJNM2014.0127

Mortazavi, F., Mehrabadi, M., Hollins Martin, C. J., & Martin, C. R. (2021). Psychometric properties of the birth satisfaction scale-revised (BSS-R) in a sample of postpartum Iranian women. *Health Care for Women International*, *42*(4-6), 836-851. https://doi.org/10.1080/07399332.2020.1802464

Nahaee, J., Mohammad-Alizadeh-Charandabi, S., Abbas-Alizadeh, F., Martin, C. R., Hollins Martin, C. J., Mirghafourvand, M., & Hassankhani, H. (2020). Pre-and during-labour predictors of low birth satisfaction among Iranian women: A prospective analytical study. *BMC Pregnancy and Childbirth*, *20*, 408. https://doi.org/10.1186/s12884-020-03105-5

Nakic Rados, S., Matijas, M., Brekalo, M., Hollins Martin, C. J., & Martin, C. R. (2023). Further validation of the birth satisfaction scale-revised: Factor structure, validity, and reliability. *Current Psychology*, *42*(16), 13693-13702. https://doi.org/10.1007/s12144-021-02688-2

Nilver, H., Begley, C., & Berg, M. (2017). Measuring women's childbirth experiences: A systematic review for identification and analysis of validated instruments. *BMC Pregnancy and Childbirth*, *17*(1), 1-19. https://doi.org/10.1186/s12884-017-1356-y

R Core Team. (2022). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. http://www.R-project.org/

Ratislavová, K., Hendrych Lorenzová, E., Hollins Martin, C. J., & Martin, C. R. (2024). Translation and validation of the Czech Republic version of the Birth Satisfaction Scale-Revised (BSS-R). *Journal of Reproductive and Infant Psychology*, *42*(1), 78-94. https://doi.org/10.1080/02646838.2022.2067837

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

Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*, 1-36. https://doi.org/10.18637/jss.v048.i02

Saudi Vision. (2023). *Kingdom of Saudi Arabia: vision 2030*. https://www.vision2030.gov.sa/en

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*(1), 108. https://doi.org/10.1186/1471-2393-13-108

Scamell, M., Altaweli, R., & McCourt, C. (2017). Sarah's birth. How the medicalisation of childbirth may be shaped in different settings: Vignette from a study of routine intervention in Jeddah, Saudi Arabia. *Women and Birth*, *30*(1), e39-e45. https://doi.org/10.1016/j.wombi.2016.08.002

Shorey, S., Yang, Y. Y., & Ang, E. (2018). The impact of negative childbirth experience on future reproductive decisions: A quantitative systematic review. *Journal of Advanced Nursing*, *74*(6), 1236-1244. https://doi.org/10.1111/jan.13534

Skvirsky, V., Taubman–Ben-Ari, O., Hollins Martin, C. J., & Martin, C. R. (2020). Validation of the Hebrew Birth Satisfaction Scale–Revised (BSS-R) and its relationship to perceived traumatic labour. *Journal of Reproductive and Infant Psychology*, *38*(2), 214-220. https://doi.org/10.1080/02646838.2019.1600666

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

Størksen, H. T., Garthus-Niegel, S., Vangen, S., & Eberhard-Gran, M. (2013). The impact of previous birth experiences on maternal fear of childbirth. *Acta Obstetricia et Gynecologica Scandinavica*, *92*(3), 318-324. https://doi.org/10.1111/aogs.12072

The International Consortium for Health Outcome Measurement. (2017). *Pregnancy and childbirth standard set and reference guide, version 1.0.3*. http://www.ichom.org/medical-conditions/pregnancy-and-child birth/

Tiwari, S. K., Murry, L., Joshi, P., Tallanao, T., Zined, R., Hollins Martin, C. J., & Martin, C. R. (2023). Translation and validation of the Hindi-Indian version of the Birth Satisfaction Scale-Revised. *Journal of Obstetrics and Gynaecology Research*, *49*(3), 938-945. https://doi.org/10.1111/jog.15520

Viirman, F., Hesselman, S., Wikström, A.-K., Svanberg, A. S., Skalkidou, A., Poromaa, I. S., & Wikman, A. (2022). Negative childbirth experience–what matters most? a register-based study of risk factors in three time periods during pregnancy. *Sexual & Reproductive Healthcare*, *34*, 100779. https://doi.org/10.1016/j.srhc.2022.100779

World Health Organization. (2018). *WHO recommendations: intrapartum care for a positive childbirth experience*. https://www.who.int/publications/i/item/9789241550215