A Systematic Review of Psychometric Properties of Breastfeeding Self-Efficacy Scale-Short Form (BSES-SF)

Masumeh Ghazanfarpour1, Malihe Afiat2, *Masaudeh Babakhanian3, Faezeh Sadat Akrami4, Leila Kargarfard5, Fatemeh Rajab Dizavandi6, Talat Khadivzadeh7

1Department of Midwifery, Razi School of Nursing and Midwifery, Kerman University of Medical Sciences, Kerman, Iran. 2Assistant Professor, Infertility Fellowship, Women's Health Research Center, Obstetrics and Gynecology Department of Obstetrics and Gynecology, Faculty of Medicine, Mashhad University of Medical Sciences, Mashhad, Iran. 3Student Research Committee, Psychiatry and Behavioral Research Centre, Mazandaran University of Medical Sciences, Sari, Mazandaran, Iran. 4Department of Clinical Psychology, School of Humanities and Social Sciences, Islamic Azad University, Science and Research Branch, Tehran, Iran. 5Instructor of Fatemeh School Nursing and Midwifery, Shiraz University of Medical Sciences, Iran. 6Department of Medical Education, Faculty of Medicine, Mashhad University of Medical Sciences, Mashhad, Iran. 7Evidence-Based Care Research Center, Mashhad University of Medical Sciences, Mashhad, Iran.

Abstract

Background: The breastfeeding self-efficacy is one of the psychological factors affecting the duration of breastfeeding; the Breastfeeding self-efficacy measured by the breastfeeding self-efficacy scale-short form (BSES-SF). It was validated and reliable in several studies. The current systematic review aimed to comprehensively review the validity and reliability of the BSES-SF in various versions to give comprehensive information for authorities in this field.

Materials and Methods: We searched English databases, including Medline (via PubMed), Scopus, Cochrane library and Web of Science since inception to May 2018 to find all previously published reports the psychometric dimensions of the Breastfeeding Self-Efficacy. Two reviewers read independently extracted data assessed. The quality of the studies was carried out by the Consensus-based standards for the selection of health status Measurement instruments (COSMIN) checklist.

Results: In this systematic review, reliability was measured using Cronbach’s alpha; inter coefficient of correlation (ICC) as well as an analysis of the item-total correlation. Cronbach’s alpha of breastfeeding self-efficacy for whole scale ranged from 0.76 to 0.95 considered moderate to excellent. Only two studies assessed test-retest reliability. One study reported correlation between two time intervals that was high (r=0.94) and another study measured test-retest reliability using ICC. It was ranged from 0.69 to 0.78 that consider suitable. Item-total correlation ranged from 0.23 to 0.85. Construct validity was tested using factor analysis. Original one-factor structure was confirmed in Spanish, Chinese, American, Canadian, Brazilian and Creation version. Some studies reported four and six –factor structure. Confirmatory factor analysis (CFA) confirmed one-factor structure.

Conclusion: Original one-factor structure was confirmed in Spanish, Chinese, American, Canadian, Swedish, and Japanese, creation, Turkish, Portuguese, Brazilian and Creation; version six and four-factor need to further test in future studies. This study provides evidence also that the modified BSES-SF may be a valid and reliable measure of breastfeeding self-efficacy.

Key Words: Breastfeeding self-efficacy, BSES-SF, Psychometric properties, Systematic review.


*Corresponding Author:
Masaudeh Babakhanian, Psychiatry and Behavioral Research Centre, Mazandaran University of Medical Sciences, Sari, Mazandaran, Iran.
Email: babakhanian.m@gmail.com

Received date: Mar.22, 2018; Accepted date: Jul.12, 2018
1- INTRODUCTION

Breast milk is the most suitable feed for infant and is the most important and effective way to maintain and sustain the health of infants. The World Health Organization and the United Nations Children's Fund have recommended that each child should have exclusive breastfeeding up to six months and continue by two years of old or more (1). Nutrition value of breast milk and its role have been proven in reducing infant mortality and increasing intelligence quotient, as well as providing the best pattern of growth and development of children and their health in adulthood (2-4). One of ways to reach to the goals of global nutritional policies, increase the level of lactation, is to assess factors effecting on breastfeeding.

Factors like socioeconomic status are less variables and healthcare providers for improving the outcome need to consider variable parameters such as maternal knowledge of the benefits of lactation, supportive systems and breastfeeding self-efficacy (5-9). The breastfeeding self-efficacy is one of the psychological factors affecting the duration of breastfeeding. The breastfeeding self-efficacy is defined as the mother's confidence in her ability to breast-feed her new infant, which is one of the factors influencing breastfeeding during the six months after delivery (2).

This self-efficacy is influenced by several factors such as mother's education, social support, type of delivery, satisfaction with childbirth analgesia, satisfaction with postpartum care, mother's understanding of lactation, the method of mother for feeding and mother's anxiety. Additionally, the psychosocial status of the mother, the support of the woman by her mother, lactation at the first hours of birth, the employment status and parity of the mother are also other items effective in breastfeeding self-efficacy (5).

The rate of exclusive breastfeeding at the international level was 56.8% and 27.7% in the fourth month and the sixth month, respectively. This is in the rural areas of Iran, 58% and 29% in the fourth month and in the sixth month, respectively (1). One-third of Canadian women have breastfeeding and their children can be ablactated eight weeks after childbirth (2). Regarding the global importance of exclusive breastfeeding and statistics, and given that the breastfeeding self-efficacy is one of the factors affecting the continuation of lactation, and both descriptive and interventional studies are ongoing in this field and are in seeking a way to promote breastfeeding.

There is a need to a question to assess breastfeeding self-efficacy in mothers. The breastfeeding self-efficacy scale-short form (BSES-SF) is self-report instrument with 33 items. It designed and developed by Dennis and Faux (10) in 1999 to evaluate mother’s confidence in breastfeeding. All items are preceded by the phrase "I can always". It scored on a 5-point Likert scale showing from 1 = not at all confident to 5 = always confident. All items are positively word and total score ranged from 33 to 165 (11). Due to the consistently high of Cronbach’s alpha values, it was decided to redundant items. In order to address to this concern, Dennis et al. in 2003 decreased items 33 to 14 (11). Short form of BSES was validated in different versions including Portuguese (12), Polish (13), Japanese (14), Spanish (15), Swedish (15), Turkish (16), and Chinese (17), and Persian. The current systematic review aimed to comprehensively review the validity and reliability of the breastfeeding self-efficacy questionnaire in various versions to give comprehensive information for authorities in this field.

2- MATERIALS AND METHODS
2-1. Method
We searched English databases, including Medline (via PubMed), Scopus, Cochran library and Web of Science since inception to May 2018 to find all previously published reports on studying the psychometric dimensions of the Breastfeeding Self-Efficacy using the keywords of (Breastfeeding Self-Efficacy Scale-Short Form or BSES-SF OR Breastfeeding) AND (Reliability OR Validity OR Psychometrics OR Validity OR reliability OR Factor Analysis OR exploratory factor analysis OR Confirmatory factor analysis OR CFA OR EFA OR Cronbach's alpha OR Test-Retest Reliability). In our review, there was no language or date restriction. We also checked the bibliographies of related articles to detect any studies not retrieved via the above mentioned electronic databases.

2-2. Data extraction
Two reviewers read independently the abstract and titles of all obtained articles, and then this selected articles to be read in details. They also separately extracted the following data using standardized data extraction form. These was including the first author of articles, area of study, age, year, sample size, study population, type of studies, timing administration of test, type of factorial structure, type of reliability and main result.

2-3. Quality of study
The quality assessment of the studies were carried out by the COSMIN checklist (18), involving internal consistency, reliability, measurement error, content validity, structure validity, and hypothesis testing, cross cultural, criterion, responsiveness, interpretability and generalizability. Within this checklist, aspects of construct validity measured by seven questions and aspects of internal consistency assessed using 11 questions (questions are shown in below Table.1 and 2) (please see the tables in the end of paper).

3- RESULTS
In total 1,413 of articles were identified through database searching. One article was identified through other sources; 774 articles remained after removal of duplicate articles; 755 articles out of 774 articles were excluded after reading title and abstract; 19 articles were assessed in detail; seven articles were excluded due to assess other psychometric properties or being type of long-form of BSES-SF and finally 12 articles were included into systematic review. Figure.1 showed search strategy of the selection of studies. The some characteristics of 12 studies included into systematic review shown in Table.3 (please see the table and figure in the end of paper).

3-1. Reliability
In this systematic review, reliability was measured using Cronbach’s alpha, inter coefficient of correlation (ICC) as well as an analysis of the item-total correlation. In Canadian version of BSES-SF, Cronbach’s alpha in a sample of mothers with preterm infants showed a value of 0.88 for whole of scale. Also, item-total correlation ranged from 0.23 to 0.70 (19). In Portuguese version of BSES-SF, Cronbach’s alpha was 0.71 for overall of scale. Lowest value for item-total correlation was 0.54 (12). In Spanish, Japanese, Swedish and Polish version, total Cronbach’s alpha of BSES-SF was near 0.90 (13-15, 20). In Swedish version, item-total correlation was in a range 0.34 to 0.74 (20). In Turkish version, total Cronbach’s alpha of BSES-SF was 0.86. Item-total correlation ranged from 0.23 to 0.65. Test-retest reliability in a two-week interval were measured using correlation and showed a high correlation (R=0.94)(16). Total Cronbach’s alpha of BSES-SF was 0.94 in a sample of
American black women and item-total correlation ranged from 0.49 to 0.85 (21). In Canadian adolescence mothers, Cronbach’s alpha was calculated for two time points (antenatal and postnatal). For the antenatal period, total Cronbach’s alpha of version of BSES-SF was 0.844 and item-total correlation ranged from 0.34 to 0.67. At one-week post-partum, total Cronbach’s alpha was 0.93 and item-total correlation ranged from 0.51 to 0.81 (22). In Brazilian version, Cronbach’s alpha and the interclass correlation coefficient (ICC) was 0.69 and 0.78, respectively (23). In Croatian version, Cronbach’s alpha was 0.86 and all of item-total correlation was positive and fell in normal range 0.3 to 0.7 (24). In Chinese version, the Cronbach’s alpha coefficient was excellent 0.95. All item-correlations were positive and above 0.50 (17).

3-2. Construct validity

3-2-1. Exploratory factor analysis (EFA)

In Canadian version of BSES-SF, Kaiser-Meyer-Olkin (KMO) was tested to measure of sampling adequacy and correlation between variables; the KMO was 0.86. Exploratory factor analysis (EFA) was conducted using principle components analysis (PCA) without rotation. One factor solution (unidimensional structure) was yielded. However, scree plot confirmed the one-factor structure of BSES-SF. This factor had an eigenvalue of 8.85 explained 47.1% of total variance. All factor loadings were above 0.3 (19). In Portuguese version of BSES-SF, EFA was conducted using maximum likelihood factor analysis. EFA identified a 6-factors structure. These factors explained 71.5% of total variance. However, scree plot, maximum likelihood, without rotation confirmed the one-factor structure of BSES-SF. The one-factor solution counted 20% of the variance (12). In Swedish version, the KMO test was 0.9. Bartlet’s test of sphericity was significant. EFA yielded one-factor structure confirmed by scree plot (20). In Spanish version, EFA using Principal Component Analysis (PCA) showed a one-factor structure. Also, all factors loading were more than 0.3 (15). In Turkish version, KMO was 0.79 and Bartlett’s test was significant (p<0.001). EFA conducted using PCA extracted four factors. Because a lots of items had cross loading, authors reported it seem that Turkish version should be one factore structure. This structure explained 30% of total variance (16). In American version, in a sample of American black women, EFA using PCA was performed and showed a single factor structure (eight values more than 1). Factor loading ranged from 0.66 to 0.87 and communality ranged from 0.36 to 0.82 (21).

In Canadian adolescents sample, KMO was 0.80, EFA using PCA without any rotation showed only one factor accounted 53% of the variance (eigenvalue of 7.50). Also, screen plot confirmed a one factor structure. Factor loading ranged from 0.61 to 0.85 (22). In Brazilian version, KMO was 0.74 and Bartlett’s Test of sphericity was significant (p<0.001). EFA using PCA without any rotation showed a four-factor solution accounted 50% of the variance. In Croatian version, EFA using maximum likelihood with rotation varimax identified one factor solution explained 33% of variance (24).

3-2-2. Confirmatory factor analysis (CFA)

Confirmatory factor analysis (CFA) was conducted on a sample of Chinese mother. According to original version, 14 items loaded on one factor. Because the assumption of normality multivariate was violated, robust maximum likelihood factor analysis (MLE) was employ to estimate parameters. The goodness of fit indices showed that model was reasonably good fit to the data. The Chi-square model was 216.6 (df=77 and p<0.001); adjusted
goodness-of-fit index (AGFI=0.99), comparative fit index (CFI=0.98); non-normed fit index (NNFI=0.97), root mean square root of approximation (RMSEA=0.099); and standardized root mean square residual (SRMR=0.05) (17).

4- DISCUSSION

Breast milk is the most complete source of nutrients and also could reduce infant mortality (1-4). It is highly recommended that each child should have exclusive breastfeeding up to six months and continue by two years of old or more (1, 25-27). Regarding the global importance of exclusive breastfeeding and statistics, and given that the breastfeeding self-efficacy is one of the factors affecting the continuation of lactation. The BSES-SF was developed to investigate mothers’ confidence in their ability to breastfeed their newborn infants. The short form of BSES-SE is a 14-item (11). The current systematic review aimed to comprehensively review the validity and reliability of the breastfeeding self-efficacy questionnaire in various versions to give comprehensive information for authorities in this field. In this systematic review, reliability was measured using Cronbach’s alpha; inter coefficient of correlation (ICC) as well as an analysis of the item-total correlation. Only two studies assessed test-retest reliability. One study reported correlation between two time intervals that was high (r=0.96), and another study measured test-retest reliability using ICC. ICC ranged from 0.69 to 0.78. Item-total correlation ranged from 0.23 to 0.85. The Cronbach’s alpha of breastfeeding self-efficacy for whole scale ranged from 0.76 to 0.95 considered moderate to excellent. The finding of systematic review showed that BSES-SF has excellent reliability Construct validity was evaluated by testing factor structure of BSES-SE. Almost all of studies employed exploratory Factor analysis (EFA) to assess factor structure. Only one study conducted confirmatory factor analysis (CFA). Original one-structure factor comforted in Spanish, Chinese, American, Canadian, Swedish, and Japanese, creation, Turkish, Portuguese, Brazilian and creation version. CFA confirmed one-factor structure. Difference in factor structures may be reflected from difference in socio-demographic, cultural and as well as methodological difference such as difference in methods of translations and difference in statistical method conducted to extract factors for example (maximum likelihood vs principle component analysis). Breastfeeding self-efficacy showed a significant relationship with maternal age, marital status (having partner or being single), previous experience to breastfeed and previous breastfeeding duration (12, 13, 23). Therefore, further studies need to be performed to assess measurement invariance in these groups. Pervious one-factor solution replicated in various populations. However, there is need for further research on two or four-factor structure in various populations.

4-1. Clinical Practice and Further Research

This instrument could be employed in future research to detect mothers are at risk of breastfeeding failure. Therefore, it can used to identify effecting factors on breastfeeding self-efficacy.

4-2. Limitation and suggestions for future

Several limitations were seen in systematic review that it is worthless to be address. Different method such as Monte Carlo and rule of thumb was used to determine sample size for exploratory factor analysis. Basis of rule of thumb, at least five cases for each item need to perform an exploratory factor analysis. Some sample sizes included into systematic review seems to be relatively small; for example
size sample was 103 in Dennis et al.’s study (22), 135 subjects in Oliver-Roig et al.’s study (15), 144 subjects in Wheeler et al.’s study (19), and 153 subjects in Deborah et al.’s study (21). KMO and Bartlett’s test measure to assess adequacy of sample size and correlation between items and it is necessary to be conduct before factor analysis. However, some studies did not report these indices. According to COSMIN checklist, percentage of missing items and how research team decides to handle of missing items should be reported. This important item was reported only in one study. The convenience sampling, instead of random sampling, was applied in all studies included into systematic review. The convenience sample as well as the use of a single site and university site limited generalizability of the finding to other environment. Further studies should focus on longitudinal data, random sampling, and multiple sitting. One important limitation is use of self-report questionnaires for data collection, may increase response bias (28).

5- CONCLUSIONS

Original one-factor structure was confirmed in in Spanish, Chinese, American, Canadian, Swedish, and Japanese, Turkish, Portuguese, Brazilian and Creation version. Six and four-factor structure need to further test in future studies. This study provides evidence also that the modified BSES-SF (short form) may be a valid and reliable measure of breastfeeding self-efficacy in various populations.

6- CONFLICT OF INTEREST: None.

7- REFERENCES


Table 1: The COSMIN checklist (aspects of internal consistency) for methodological quality of studies (11).

<table>
<thead>
<tr>
<th>Authors, Reference</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deborah et al., (21)</td>
<td>Unclear</td>
<td>No</td>
<td>No</td>
<td>Unclear</td>
<td>Yes</td>
<td>Yes</td>
<td>Not applicable</td>
<td>No</td>
<td>Yes</td>
<td>Not applicable</td>
<td>No</td>
</tr>
<tr>
<td>Dennis et al., (22)</td>
<td>Unclear</td>
<td>No</td>
<td>No</td>
<td>Unclear</td>
<td>Yes</td>
<td>Unclear</td>
<td>Not applicable</td>
<td>No</td>
<td>Yes</td>
<td>Not applicable</td>
<td>No</td>
</tr>
<tr>
<td>Dodt et al., (23)</td>
<td>Unclear</td>
<td>No</td>
<td>No</td>
<td>Unclear</td>
<td>Yes</td>
<td>Unclear</td>
<td>Not applicable</td>
<td>No</td>
<td>Yes</td>
<td>Not applicable</td>
<td>No</td>
</tr>
<tr>
<td>Bosnjak et al., (24)</td>
<td>Unclear</td>
<td>No</td>
<td>No</td>
<td>Unclear</td>
<td>Yes</td>
<td>Unclear</td>
<td>Not applicable</td>
<td>No</td>
<td>Yes</td>
<td>Not applicable</td>
<td>No</td>
</tr>
<tr>
<td>Ip et al., (17)</td>
<td>Unclear</td>
<td>No</td>
<td>No</td>
<td>Unclear</td>
<td>Yes</td>
<td>Unclear</td>
<td>Not applicable</td>
<td>No</td>
<td>Yes</td>
<td>Not applicable</td>
<td>No</td>
</tr>
<tr>
<td>Oliver-Roig et al., (15)</td>
<td>Unclear</td>
<td>No</td>
<td>No</td>
<td>Unclear</td>
<td>Yes</td>
<td>Unclear</td>
<td>Not applicable</td>
<td>No</td>
<td>Yes</td>
<td>Not applicable</td>
<td>No</td>
</tr>
<tr>
<td>Wheeler et al., (19)</td>
<td>Unclear</td>
<td>Yes</td>
<td>Yes</td>
<td>Unclear</td>
<td>Yes</td>
<td>Unclear</td>
<td>Not applicable</td>
<td>No</td>
<td>Yes</td>
<td>Not applicable</td>
<td>No</td>
</tr>
<tr>
<td>Aydin et al., (16)</td>
<td>Unclear</td>
<td>No</td>
<td>No</td>
<td>Unclear</td>
<td>Yes</td>
<td>Unclear</td>
<td>Not applicable</td>
<td>No</td>
<td>Yes</td>
<td>Not applicable</td>
<td>No</td>
</tr>
<tr>
<td>Zubaran et al., (12)</td>
<td>Unclear</td>
<td>No</td>
<td>No</td>
<td>Unclear</td>
<td>Yes</td>
<td>Unclear</td>
<td>Not applicable</td>
<td>No</td>
<td>Yes</td>
<td>Not applicable</td>
<td>Yes</td>
</tr>
<tr>
<td>Nanishi et al., (14)</td>
<td>Unclear</td>
<td>No</td>
<td>No</td>
<td>Unclear</td>
<td>Yes</td>
<td>Unclear</td>
<td>Not applicable</td>
<td>No</td>
<td>Yes</td>
<td>Not applicable</td>
<td>Yes</td>
</tr>
<tr>
<td>Wutke et al., (13)</td>
<td>Unclear</td>
<td>No</td>
<td>No</td>
<td>Unclear</td>
<td>Yes</td>
<td>Unclear</td>
<td>Not applicable</td>
<td>No</td>
<td>Yes</td>
<td>Not applicable</td>
<td>No</td>
</tr>
<tr>
<td>Gerhardsson et al., (20)</td>
<td>Unclear</td>
<td>No</td>
<td>No</td>
<td>Unclear</td>
<td>Yes</td>
<td>Unclear</td>
<td>Not applicable</td>
<td>No</td>
<td>Yes</td>
<td>Not applicable</td>
<td>No</td>
</tr>
</tbody>
</table>

1. Do the scales have effect indicators, meaning is a reflective model a basis?  
2. Design requirements?  
3. Was the percentage of missing items presented?  
4. What was a description of how missing items were addressed?  
5. Was there adequacy for the sample size in the internal consistency analysis?  
6. Was the unidimensionality of the scale evaluated, meaning was factor analysis or IRT model used?  
7. Was there adequacy for the sample size in the unidimensionality analysis?  
8. Was there separate internal consistency statistic calculated for each unidimensional subscale?
9. Were there any key flaws in the study design or methods?
10. Methods of statistical analysis
11. Was Cronbach’s alpha obtained for Classical Test Theory (CTT)?
12. Was Cronbach’s alpha or KR-20 obtained for dichotomous scores?
13. Was a goodness of fit statistic at a global level obtained for IRT? e.g. Chi-squared test.

IRT; Item Response Theory; CTT: Classical Test Theory.

<table>
<thead>
<tr>
<th>Authors, Reference</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deborah et al., (21)</td>
<td>Unclear</td>
<td>No</td>
<td>No</td>
<td>Unclear</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Dennis et al., (22)</td>
<td>Unclear</td>
<td>No</td>
<td>No</td>
<td>Relatively Small</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Dodt et al., (23)</td>
<td>Unclear</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Bosnjak et al., (24)</td>
<td>Unclear</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>IP et al., (17)</td>
<td>Unclear</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Oliver-Roig et al., (15)</td>
<td>Unclear</td>
<td>No</td>
<td>No</td>
<td>Relatively Small</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Wheeler et al., (19)</td>
<td>Unclear</td>
<td>Yes</td>
<td>Yes</td>
<td>Relatively Small</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Aydin et al., (16)</td>
<td>Unclear</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Gerhardsson et al., (20)</td>
<td>Unclear</td>
<td>Yes</td>
<td>Yes</td>
<td>Relatively Small</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Zubaran et al., (12)</td>
<td>Unclear</td>
<td>Yes</td>
<td>Yes</td>
<td>Relatively Small</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Nanishi et al., (14)</td>
<td>Unclear</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Gerhardsson et al., (20)</td>
<td>Unclear</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
</tbody>
</table>

1. Do the scales have effect indicators, meaning is a reflective model a basis?
2. Was the percentage of missing items presented?
3. What was a description of how missing items were addressed?
4. Was there adequacy for the sample size in the analysis?
5. Were there any key flaws in the study design or methods?
6. Was exploratory or confirmatory factor analysis carried out for CTT?
7. Were IRT tests carried out for IRT to detect the unidimensionality of the items?

IRT; Item Response Theory; CTT: Classical Test Theory.
Fig. 1: PRISMA flowchart of study.
**Table 3: The characteristic of 12 studies included into systematic review.**

<table>
<thead>
<tr>
<th>Authors, Area of study, Year, Reference</th>
<th>Age, year</th>
<th>Sample size</th>
<th>Study population</th>
<th>Type of studies</th>
<th>Timing administration of test</th>
<th>Factorial structure</th>
<th>Reliability</th>
<th>Main result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deborah et al., (21), 2010, USA</td>
<td>18 or older</td>
<td>153 mother</td>
<td>Black Women in the USA</td>
<td>NM</td>
<td>4-week postpartum</td>
<td>EFA</td>
<td>Internal consistency and item-total correlation</td>
<td>Total cronbach’s alpha was 0.94 and item-total correlation ranged from 0.49 to 0.85. EFA identified single factor structure Factor with factor loading ranged from 0.66 to 0.87.</td>
</tr>
<tr>
<td>Dennis et al., (22), 2011, Canada</td>
<td>16.82 ± 1.18</td>
<td>103 adolescents</td>
<td>Canadian adolescents</td>
<td>NM</td>
<td>Antenatal and postnatal</td>
<td>EFA</td>
<td>Internal consistency and item-total correlation</td>
<td>Cronbach’s alpha was 0.844 and 0.93 antenatal and post-natal, respectively. Item-total correlation ranged from 0.34 to 0.67 for the antenatal assessment and from 0.51 to 0.81 for posts natal assessment. EFA extracted only one factor structure accounted 53% of the variance.</td>
</tr>
<tr>
<td>Dodt et al., (23), 2012, Brazil</td>
<td>23.9</td>
<td>294 postpartum women</td>
<td>Brazilian women with low income</td>
<td>NM</td>
<td>Postpartum</td>
<td>EFA</td>
<td>Internal consistency and item-total correlation</td>
<td>In Brazilian, cronbach’s alph was 0.74 and 078% of item-total correlation failed in normal rang 0.3 to 0.7. Interclass correlation coefficient (ICC) ranged from 0.69 and 0.78, EFA using PCA without any rotation showed only one factor accounted 57% of the variance.</td>
</tr>
<tr>
<td>Pavlicic Bosnjak et al., (24), 2012, Croatia</td>
<td>30 (± 4)</td>
<td>190 breastfeeding mothers</td>
<td>Croatian mothers</td>
<td>NM</td>
<td>1 and 6 postpartum</td>
<td>EFA</td>
<td>Internal consistency and item-total correlation</td>
<td>Cronbach’s alph was 0.86 and all of item-total correlation failed in normal rang 0.3 to 0.7. EFA using maximum likelihood with rotation varimax identified one factor solution explained 33 % of variance.</td>
</tr>
<tr>
<td>Ip et al., (17), 2012, China</td>
<td>30</td>
<td>185 breastfeeding mothers</td>
<td>Chinese mothers</td>
<td>NM</td>
<td>1 and 6 postpartum</td>
<td>CFA</td>
<td>Internal consistency and item-total correlation</td>
<td>The Cronbach’s alpha coefficient was excellent 0.95. All item-correlations were positive and above 0.50. The goodness of fit indices showed that model was resonably good fit to the data.</td>
</tr>
<tr>
<td>Oliver-Roig (15), 2012, Spain</td>
<td>30.1 ± 55.03</td>
<td>135 breastfeeding mothers</td>
<td>Spanish mothers</td>
<td>Methodological study</td>
<td>The second day postpartum</td>
<td>EFA</td>
<td>Internal consistency</td>
<td>Total cronbach’s alpha of of BSES-SF was 0.92 . EFA using PCA showed a one factor sloution Also all factor loading were more than 0.3.</td>
</tr>
<tr>
<td>Wheeler et al. (19), 2012, Canada</td>
<td>29.8 ± 5.2</td>
<td>144 breastfeeding mothers</td>
<td>Canadian mothers with preterm infants</td>
<td>Methodological study</td>
<td>I-week post partum</td>
<td>EFA</td>
<td>Internal consistency and item-total correlation</td>
<td>Total cronbach’s alpha of BSES-SF was 0.88. item-total correlation ranged from 0.23 to 0.70. EFA using PCA without rotation showed one factor sloution explained 47.1% of total variance. All factor loading was above 0.3.</td>
</tr>
</tbody>
</table>
### Psychometric Properties of BSES-SF

<table>
<thead>
<tr>
<th>Authors</th>
<th>Methodological Study</th>
<th>Internal Consistency</th>
<th>Item-total Correlation</th>
<th>EFA</th>
<th>Test-retest Reliability</th>
<th>Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aydin et al., 2018, Turkey</td>
<td>30% between 19-24, 33.7% between 25-29%, 36% between 30-34</td>
<td>0.94</td>
<td>0.23 to 0.65</td>
<td>EFA</td>
<td>Total Cronbach’s alpha was 0.86. Item-total correlation ranged from 0.23 to 0.65. EFA extracted four factors. Because a lot of items had cross loading, authors reported it seems that Turkish version should be one factor structure.</td>
<td></td>
</tr>
<tr>
<td>Zabaran et al., 2010, Portugal</td>
<td>25.4</td>
<td>89 mothers in Southern Brazil</td>
<td>NM</td>
<td>Postpartum</td>
<td>EFA</td>
<td>Total Cronbach’s alpha was 0.71. Item-total correlation ranged from 0.23 to 0.70. EFA identified 6 factors explained 71.5% of total variance. However, scree plot suggested the one-factor structure counted 20% of the variance.</td>
</tr>
<tr>
<td>Wutke et al., 2006, Poland</td>
<td>28</td>
<td>104 mothers</td>
<td>Poland mothers</td>
<td>Methodological study</td>
<td>6-16 postpartum</td>
<td>EFA</td>
</tr>
<tr>
<td>Nanishi et al., 2015, Japan</td>
<td>30.8±4.8</td>
<td>378 women in their third trimester</td>
<td>Japanese mothers</td>
<td>A secondary analysis of intervention study</td>
<td>Third trimester of pregnancy</td>
<td>EFA</td>
</tr>
<tr>
<td>Gerhardsson et al., 2014, Swedish</td>
<td>Median=31</td>
<td>120 mothers</td>
<td>Swedish mother</td>
<td>NM</td>
<td>During the first week postpartum</td>
<td>EFA</td>
</tr>
</tbody>
</table>