# Breastfeeding Self-Efficacy: A Systematic Review of Psychometric Properties Using COSMIN


## Abstract

**Background:** Breastfeeding self-efficacy has been proven to play a predictive role in enhancing breastfeeding initiation and continuation. Breastfeeding self-efficacy measurement tools have facilitated healthcare professionals' early identification and support of women at higher risk of early discontinuation of breastfeeding.

**Research aim:** The aim of this study was to assess the psychometric properties of breastfeeding self-efficacy measurement tools.

**Method:** A systematic review was carried out in three phases. Phase one comprises a systematic literature review performed in PubMed, SCOPUS, Web of Science, and Cochrane Database of Systematic Reviews from February 2021 to January 2023. The included articles were 36 studies. Phase two provided a quality appraisal of the psychometric properties of each of the seven breastfeeding self-efficacy measurement tools, according to COnsensus-based Standards for the selection of health Measurement Instrument checklist (COSMIN) guidelines. Phase three summarized and graded the overall quality of evidence using the Grading of Recommendations Assessment, Development and Evaluation (GRADE) modified approach.

**Results:** The included articles involved 9,225 women and encompassed 7 breastfeeding self-efficacy measurement tools. The Breastfeeding Self-Efficacy Scale, Breastfeeding Self-Efficacy Scale – Short Form (BSES-SF), and Prenatal Breastfeeding Self-Efficacy Scale were supported by Grade A evidence sustaining their validity and reliability to assess breastfeeding self-efficacy in the continuum of maternity care. The BSES-SF is the most feasible tool in clinical practice and the most utilized internationally available in fifteen languages.

**Conclusion:** This systematic review provided a Grade A recommendation on breastfeeding measurement tools that will be helpful both for clinical and research purposes.

Registered in the International Prospective Register of Systematic Reviews (PROSPERO; CRD42021238450)

## Background

Breastfeeding is one of the most effective practices to promote child survival and wellbeing and improve mother's physical and mental health (UNICEF, 2018). However, the exclusive breastfeeding rate worldwide is low at 44% in the first six months (WHO, 2021), and the World Health Organization (WHO) has included the target of 50% by 2025 in the global nutrition targets (WHO/UNICEF, 2019). In this global context, understanding the factors that affect the promotion of breastfeeding practice is a matter of public health. Several studies have described modifiable factors associated with the early discontinuation of breastfeeding practice. Breastfeeding self-efficacy can be defined as a woman's confidence in her ability to breastfeeding, and it is often associated with breastfeeding satisfaction; both factors have been recognized to play a pivotal role in undermining optimal behaviors associated with breastfeeding (De Roza et al., 2019). Furthermore, self-efficacy is a predictor of an effective initiation of breastfeeding and its duration (Tuthill et al., 2016).

Self-efficacy was conceptualized for the first time by Albert Bandura in social cognitive theory and defined as the individual's belief to succeed in a specific situation (Bandura, 1977). According to Bandura, self-efficacy is crucial in determining a specific behaviour since it reflects individuals' perceptions about their abilities, which are linked to contextual and specific situations (Scholz et al., 2002). Within Bandura's framework, Dennis et al. (1999) developed a breastfeeding self-efficacy theory, where breastfeeding self-efficacy refers to a mother's confidence in her perceived ability to breastfeed the baby (Dennis, 1999). In other words, mothers with a higher self-efficacy are more likely to breastfeed infants, persist over their daily challenges, and react positively in dealing with the difficulties regarding breastfeeding (Dennis, 1999).

Consistent with Bandura's theory, Dennis stated that four main sources could influence breastfeeding self-efficacy: past experiences, vicarious experiences, verbal persuasion, and

physiological responses (Bandura, 1977, 2001; Dennis, 1999). Past experiences are mainly based on reproducing a sense of personal mastery of a challenging situation when the mother experienced winning behaviours over difficulties. Vicarious experience is based on acquiring experience through social comparison. In this regard, verbal persuasion and physiological responses are the reactions to the verbal encouragement of peers that produce emotional arousal, which is the sensation of being able to handle a difficult circumstance (Bandura et al., 1999; Dennis, 1999). In this regard, verbal persuasion refers to the act of receiving positive feedback or verbal encouragement from others, such as peers or authority figures, that can boost an individual's confidence in their ability to perform a specific task. This feedback can come in the form of praise, reassurance, or motivational statements and can significantly affect an individual's self-efficacy. On the other hand, physiological responses refer to the bodily sensations that occur in response to verbal persuasion. These sensations may include an increase in heart rate, breathing rate, or sweating and are often linked to emotional arousal. When individual experiences positive physiological responses, it can be an indication that they are feeling more confident in their ability to handle a difficult situation, which can further boost their self-efficacy.

Interventional studies that utilized the theory around breastfeeding self-efficacy were recently summarized in a Cochrane review that included 116 trials with over 98,000 mother-infant pairs and found that the support for exclusive breastfeeding, including the enhancement of self-efficacy levels, is likely to reduce the number of women stopping breastfeeding, particularly in the first few months (Gavine et al., 2022). For this reason, healthcare professionals might use these four sources to enhance breastfeeding self-efficacy to plan specific educational and motivational activities, which might be based on precise self-efficacy assessments (Dennis, 1999). Several self-report tools were developed in the last two decades to assess breastfeeding self-efficacy in heterogeneous social and cultural contexts and different populations, both antenatally and postnatally, which have been described and critically appraised in a recent systematic review (Tuthill et al., 2016). These self-report tools encompass different items with a few differences in the

specific measured domains and the number of items per scale, even if all the tools aim to measure the same theoretical construct, namely breastfeeding self-efficacy. The advantages of adopting these assessments in practice are related to the early identification of women at higher risk of interrupting breastfeeding (Dennis, 1999; Gavine et al., 2022). Early identification of risks allows healthcare professionals to support women by employing a strategy to enhance breastfeeding self-efficacy (Gavine et al., 2022, pp. 4-8).

Although self-report tools for breastfeeding self-efficacy have been published in several contexts, the quality and quantity of the available evidence sustaining the validity and reliability of each available tool are heterogeneous and, thus far, not yet synthesized (Tuthill et al., 2016). A synthesis of the quality measurement properties of the available breastfeeding self-efficacy tools could help healthcare professionals accurately measure breastfeeding self-efficacy and, therefore, plan specific and individualized educational and supportive strategies. Furthermore, a synthesis of the quality measurement properties of these tools is required to provide recommendations on the most suitable tool to measure breastfeeding self-efficacy by accounting for contexts, cultural, and clinical elements (Mokkink, Prinsen, et al., 2018). For this reason, this study was aimed to assess the psychometric properties of breastfeeding self-efficacy measurement tools.

# Method

## Research Design

This systematic review was designed using the COnsensus-based Standards for the selection of health Measurement Instrument (COSMIN) methodology (Prinsen et al., 2018). The rationale for using this design is that the COSMIN methodology allows for a more rigorous and standardized evaluation of Patient-Reported Outcome Measures (PROMs), which can help to improve the quality and reliability of PROMs in research and clinical practice (Prinsen et al., 2018).

## Sample

The target population was represented by women in a physiological maternal-neonatal dyad condition, and the literature topic to be reviewed was the validation studies of self-report

breastfeeding self-efficacy tools based on the conceptual framework of Bandura (or Dennis in the specific field of breastfeeding) in a physiological maternal-neonatal dyad condition.

The inclusion criteria were (a) papers in any language accessible to the authors (i.e., English and Italian) and published after the first breastfeeding self-efficacy tool developed by Dennis et al. in 1999; (b) validation studies of self-report breastfeeding self-efficacy tools based on the conceptual framework of Bandura (or Dennis in the specific field of breastfeeding); (c) that consider breastfeeding in a physiological maternal-neonatal dyad condition (Bandura et al., 1999; Bandura, 1977; Dennis, 1999). The exclusion criteria were given by contents focused on specific diseases of the infants or the mothers were excluded. In fact, we have considered women of childbearing age for all age groups and in low- or middle-income countries.

The PRISMA statement and flowchart were used for reporting the selection process (Page et al., 2021) (**Figure 1**). All eligible studies have been screened by two authors (GM; GG), and data extraction was performed using a predefined data collection form. The following data has been extracted: (a) first author, country, and year of publication; (b) sample characteristic and setting; (c) validated tool and availability of language-specific versions; (d) characteristics (e.g., factorial domains, number of items, rating score); and (e) a qualitative summary description of the psychometric proprieties' domains. The entire process was supervised by a third author (AM) that checked the accuracy of the information extracted.

More precisely, the literature search identified 845 records, of which 104 were duplicates. In the screening phase, 699 papers were excluded after reading the title and abstract. Therefore, 42 articles were assessed for the eligibility phase, and six papers were excluded after carefully reading the full texts. Among these, two studies were excluded due to involved ill babies and fathers' group (Dennis et al., 2018; Wheeler & Dennis, 2013), and two other studies did not consider the self-efficacy theory as a conceptual framework (Kronborg & Væth, 2019; Palmér & Jutengren, 2019), and the last two were not validation studies (Gerçek et al., 2017; Ingram et al., 2015). Thus, 36 studies enrolling 9,225 women were included describing seven breastfeeding self-efficacy tools:

Breastfeeding Self-Efficacy Scale (BSES) (n=six studies) (Creedy et al., 2003; Dai & Dennis, 2003; Dennis & Faux, 1999; Eksioglu & Ceber, 2011; Oriá et al., 2009; Torres et al., 2003); Breastfeeding Self-Efficacy Scale – Short Form (BSES-SF) (n=twenty-two studies) (Amini et al., 2019; Asgarian et al., 2019; Basu et al., 2020; Bosnjak et al., 2012; Brandão et al., 2018; Dennis, 2003; Dennis et al., 2011; Economou et al., 2021; Gerhardsson et al., 2014; Gregory et al., 2008; Iliadou et al., 2020; McCarter-Spaulding & Dennis, 2010; McQueen et al., 2013; Nanishi et al., 2015; Oliver-Roig et al., 2012; Petrozzi & Gagliardi, 2016; Radwan et al., 2022; Tokat et al., 2010; Wan-Yim et al., 2012, 2016; Wutke & Dennis, 2007; Zubaran et al., 2016); Prenatal Breastfeeding Self-Efficacy Scale (PBSES) (n=four studies) (Aydın & Pasinlioğlu, 2018; Hazar & Akça, 2018; Pineiro-Albero et al., 2013; Wells et al., 2006); Prenatal Breastfeeding Self-Efficacy Scale – Short Form (PBSES-SF) (n= one study) (Silva-Tubio et al., 2021); Prenatal rating of Efficacy in Preparation to Breastfeed Scale (PREPBS) (n= one study) (Mckinley et al., 2019). Breastfeeding Self-Efficacy Scale to Assess Exclusive Breastfeeding (BSES-EBF) (n= one study) (Boateng et al., 2019); Breastfeeding Personal Efficacy Beliefs Inventory (BPEBI) (n=one study) (Cleveland & McCrone, 2005). A description of included studies and participants group is reported in **Table 1**.

**Data Collection**

The research design was developed in three main phases. In phase one, a preliminary literature search strategy was conducted to identify the breastfeeding self-efficacy tools. After that, in phase two, a quality appraisal of psychometric proprieties of the breastfeeding self-efficacy tools was performed according to the COSMIN guideline (Mokkink, Prinsen, et al., 2018; Williamson et al., 2017). Finally, in the third phase (data analysis), a synthesis of evidence and a grade of recommendation was developed to help researchers and healthcare professionals identify the most suitable self-report tools to measure breastfeeding self-efficacy.

***Literature search strategy (Phase One)***

A systematic literature search strategy was conducted on the electronic databases of PubMed, SCOPUS, Web of Science (WOS), and Cochrane Database of Systematic Reviews from

February 2021 to January 2023 using the following keywords to create the research queries: *breastfeeding; self-efficacy; measurement; instrument; scale; questionnaire*. In this regard, we have used the comprehensive and sensitive Patient-Reported Outcome Measure(s) (PROM) filter for the PubMed database validated by the COSMIN group (Prinsen et al., 2018), which has also been adapted for the other databases. In addition, the reference lists of the identified studies have been checked to retrieve the most relevant studies.

***Quality assessment (Phase Two)***

In this phase, we have considered eight measurement proprieties for each self-report breastfeeding self-efficacy tool: content validity, structural validity, internal consistency, cross-cultural validity/measurement invariance, reliability, hypotheses testing for construct validity, criterion validity, and responsiveness (Elsman et al., 2022; Mokkink et al., 2006; Prinsen et al., 2018). The COSMIN manual includes two additional evaluation parameters: PROM development and measurement error. While PROM development is not a measurement property in itself, it considers the characteristics of developmental studies that often test for content validity, which is a measurement property. Therefore, the characteristics of PROM development are included in the overall evaluation of content validity. For this reason, we focused on content validity as a more comprehensive evaluation of the measurement properties of the included tools. In addition, measurement error is a measurement property that refers to the degree to which the scores of a PROM are affected by random error. According to the COSMIN manual, measurement error is meaningful when information regarding the minimal important change (MIC) or smallest detectable change (SDC) is available (Mokkink, Prinsen, et al., 2018). MIC is the smallest change in score on a PROM that is considered to be clinically meaningful, while SDC is the smallest change in score on a PROM that can be detected with a certain level of confidence or precision. SDC is also associated with the PROM's responsiveness because it includes assessing how small enough SDC has to be to detect meaningful changes over time or in response to an intervention (Terwee et al., 2018). In relation to breastfeeding self-efficacy, these aspects may not be meaningful in terms of

clinical aspects. This is because the focus on assessing self-efficacy in relation to breastfeeding is not to discriminate clinically meaningful changes but rather to determine how confident a woman is in dealing with the challenges of breastfeeding.

Content validity is a measurement property regarding the degree to which the content of the self-report tools adequately reflects the construct to be measured (Mokkink, de Vet, et al., 2018). Structural validity is defined by the dimensionality (factor structure) of each tool. Internal consistency is a measurement property of reliability that evaluates the degree of the interrelatedness among the items of the self-report tools. Cross-cultural validity enables the assessment of the degree to which the items' performance of the self-report tools translated or culturally adapted is an adequate reflection of the original tool. The reliability property assesses the degree to which the score remains the same for patients over time (i.e., test-retest reliability form) or under different conditions (i.e., intra-rater or internal consistency form). Hypothesis testing for construct validity allows researchers to assess the degree to which the score is consistent with the hypothesis: it could be evaluated considering the relationship of scores between self-report tools measuring the same construct (convergent validity) or different constructs (divergent validity) or testing differences between relevant groups (known-groups validity). Criterion validity is defined as the degree to which a PROM score adequately reflects a gold standard (Mokkink, de Vet, et al., 2018).

These measurement properties were assessed considering two main parameters: the overall methodological quality score (M) and the quality level of evidence (Q) (Mokkink et al., 2016). The methodological quality score includes the evaluation of the study's design, conduct, and analysis. Therefore, its evaluation considers the sample size, missing data, appropriateness of statistical methods, and other aspects related to the study design and methods used to assess the measurement property. The methodological quality of each measurement property was assessed using the COSMIN Risk of Bias checklist and applying updated criteria for good measurement properties (Mokkink, de Vet, et al., 2018). The COSMIN checklist is structured into nine boxes measuring different proprieties; each box includes specific standards of assessment on the study design and

statistical analysis (from a minimum of five to a maximum of 18 standards per box, based on the measurement properties). Accordingly, two authors (GB; AM) independently evaluated the methodological quality by assigning for each standard in the box a score on a four-point scoring system (V= very good; A= adequate; D= doubtful; I=inadequate; N=not applicable); any disagreements among the authors were sorted through a discussion. The overall quality score for each property was obtained by considering the lowest score awarded in the box.

Quality rating, on the other hand, refers to the extent to which the study provides evidence of the measurement property. The quality level of evidence evaluation assesses the relevance, internal consistency, reliability, measurement error, content validity, structural validity, hypotheses testing, cross-cultural validity, and responsiveness of the measurement property (Mokkink et al., 2016). In the COSMIN approach, the overall methodological quality score for each property is evaluated based on the updated criteria for good measurement properties described by Mokkink et al. (Mokkink, de Vet, et al., 2018). The evaluation using these criteria was performed by reviewing the quality of the measurement properties used in the study and assessing its adherence to the criteria. More precisely, Mokkink et al. (2018) have updated the criteria for good measurement properties based on a Delphi study. The update includes a clearer separation of measurement properties, a revised structure for evaluating the quality of measurement properties, and new criteria for assessing PROMs' responsiveness. For instance, internal consistency is assessable by employing the criterium based on determining whether Cronbach's alphas for each unidimensional scale or subscale are above or equal to 0.70, which is considered an acceptable threshold (Mokkink, de Vet, et al., 2018). In addition, highly correlated items within the same measure represent a second criterium required to achieve sufficient quality ratings for internal consistency. The same approach is applied to evaluate other measurement properties. The overall quality rating can be categorized as sufficient (+), insufficient (-), or indeterminate (?) based on the criteria specified in the COSMIN manual (Mokkink, Prinsen, et al., 2018). The rating of each measurement property provides

valuable information to users about the strengths and weaknesses of the psychometric properties of the PROM, which helps in selecting the appropriate tool for a specific research question.

***Overview of the COSMIN methodology***

This methodology was developed as the result of an international Delphi study involving 21 countries and 159 experts (Terwee et al., 2018) to update a previoyusly developed COSMIN standars for evaluating the quality of the content validity of research involving PROMs (Prinsen at al., 2016). Compared to other systematic review methods, the COSMIN methodology provides a more comprehensive and standardized approach to evaluating PROMs (Mokkink et al., 2006). Other systematic review methods may focus on a limited number of measurement properties or may not provide a standardized approach for evaluating PROMs. The COSMIN methodology is a widely recognized and highly regarded approach for evaluating the quality of PROMs. More precisely, it consists of a comprehensive framework that provides a standardized and transparent way of evaluating the methodological quality of PROMs, as per the case of breastfeeding self-efficacy measurements. The COSMIN methodology includes a checklist that evaluates the methodological quality of PROMs across several measurement properties and allows researchers to assess the quality of PROMs in a standardized and comprehensive manner (Elsman et al., 2022; Mokkink, de Vet, et al., 2018; Mokkink et al., 2006; Terwee et al., 2018).

The COSMIN checklist is a tool used to evaluate the methodological quality of studies that assess the psychometric properties of PROMs (Mokkink, de Vet, et al., 2018). It consists of 10 boxes, each representing a measurement property, such as reliability or construct validity. Each box contains a set of standards and criteria that need to be met to obtain a high-quality rating for the corresponding measurement property. The checklist can be used to evaluate the overall methodological quality of a study or to evaluate individual measurement properties separately. The COSMIN checklist is not meant to evaluate the risk of bias or the quality of reporting by employing scales of assessment (Mokkink, de Vet, et al., 2018). This is because the COSMIN checklist provides itself with a direct assessment of each measurement property, and each property has its

own set of standards and criteria that must be met. The traditional risk of bias (and quality evaluation of the reporting) scales do not take into account the specific requirements for each measurement property, and therefore, they are not meaningful or consistent with the COSMIN manual. Instead, the COSMIN checklist provides a comprehensive and standardized approach to evaluating the methodological quality of PROMs and their psychometric properties, and it is designed to assess each property separately. The COSMIN checklist (i.e., COSMIN Risk of Bias checklist) by providing a direct assessment of each measurement property, helps to ensure that the evaluation is specific and appropriate for the property being assessed, and it helps to avoid the risk of bias or quality issues that can arise from using generalized scales of assessment by employing the approach per se without the need for other reporting checklists or scales to perform a critical appraisal of the included studies (Mokkink, de Vet, et al., 2018).

**Measurement**

Information regarding the characteristics of breastfeeding tools was extracted, including the country(ies) where the included study was implemented, the language used to develop the breastfeeding tools, the number of included participants for each study (sample size), and the setting and timing of data collection.

As per the COSMIN methodology described in the data collection section, the variables characterizing the measurements of this systematic review and the process used to extract data from the included articles were described in the data collection, quality assessment (phase two), in relation to content validity, structural validity, internal consistency, cross-cultural validity/measurement invariance, reliability, hypotheses testing for construct validity, criterion validity, and responsiveness (Elsman et al., 2022; Mokkink et al., 2006; Prinsen et al., 2018).

The synthesis and recommendations in the data collection section described the employed process for categorizing the extracted variables, its rationale, and any relevant information about the process used to organize and summarize the data (phase three: data analysis).

**Data Analysis**

This is phase three of the study, providing the strategy for developing the final synthesis and recommendations. The pooled data for each breastfeeding self-efficacy tool involved qualitatively summarizing the results of each study that evaluated the measurement properties of each tool per measurement property (e.g., internal consistency, test-retest reliability). As per the COSMIN manual, we did not calculate the ratings quantitatively (e.g., as a percentage) but rather synthesized the findings narratively, comparing each study's results against the updated criteria for good measurement properties (Mokkink, Prinsen, et al., 2018). Hence, the overall rating has been assigned as sufficient [+], insufficient [-], inconsistent [+/-], or indeterminate [?]. In this scenario, we can achieve a sufficient rating when roughly 75% of the pooled data have satisfied the predefined criteria. For example, if a measurement property had eight studies assessing its validity, and six of them met the predefined criteria, the validity rating for that property would be sufficient because 75% of the studies satisfied the criteria described by Mokkink et al. (2018). This threshold was based on previous recommendations and discussions in the COSMIN community (Prinsen et al., 2018).

In the context of COSMIN, the Grading of Recommendations Assessment, Development and Evaluation (GRADE) modified approach is used to determine the overall quality of the evidence for each measurement property of the included studies (Prinsen et al., 2018). The GRADE-modified approach includes four levels of evidence quality: very low, low, moderate, and high. The GRADE-modified approach takes into account several factors, including the study design, risk of bias, inconsistency, indirectness, imprecision, and publication bias. Each measurement property is rated according to the quality of the evidence, with higher quality evidence indicating a greater confidence in the results. The GRADE-modified approach allows for a standardized and transparent assessment of the quality of evidence, which is essential for making evidence-based recommendations. The quality level of evidence can be downgraded or upgraded to three levels per factor, resulting in an overall quality rating reflecting confidence in the measurement property. If the study design for a particular measurement property is at high risk of bias, the quality of evidence

may be downgraded. For example, suppose that a study that assessed the content validity of the analyzed tool for assessing breastfeeding self-efficacy used a non-systematic approach to select experts to rate the relevance of the items without any clear criteria for their selection. This scenario would introduce a high risk of bias, as the experts may not have been the most relevant or knowledgeable for the task. In this case, the quality of evidence for the content validity of the specific tool would be downgraded (e.g., from high to moderate), as the study design for this measurement property is doubtful or inadequate.

In this regard, two authors (GB; AM) rated and graded the summarized results for each measurement property. This approach generally indicates how confident we are that the overall ratings are trustworthy for each measurement property and tool separately (Schünemann et al., 2013). Furthermore, when pooled data per specific properties were inconsistent or indeterminate, it was not possible to proceed with grading the overall quality of evidence using the GRADE-modified approach.

## Results

### Characteristics of the self-report tools

The BSES is a widely used self-report tool for measuring a mother's breastfeeding self-efficacy after birth. It has been validated in five languages and six studies were published between 1999 and 2011 (Eksioglu & Ceber, 2011; Torres et al., 2003; Dai & Dennis, 2003; Oriá et al., 2009; Creedy et al., 2003; Dennis & Faux, 1999). The BSES-SF, a short form of BSES, has been further validated in fifteen languages and tested in various groups, such as pregnant adolescents and multi-ethnic minorities (Amini et al., 2019; Asgarian et al., 2019; Wan-Yim et al., 2012, 2016; Tokat et al., 2010; Brandão et al., 2018; Zubaran et al., 2016; Gerhardsson et al., 2014; Petrozzi & Gagliardi, 2016; Nanishi et al., 2015; Oliver-Roig et al., 2012; Wutke & Dennis, 2007; Basu et al., 2020; Bosnjak et al., 2012; Economou et al., 2021; Iliadou et al., 2020; Dennis, 2003; Dennis et al., 2011; Gregory et al., 2008; McCarter-Spaulding & Dennis, 2010; McQueen et al., 2013; Radwan et al., 2022).

An adapted version of BSES-SF, called BSES-EBF, was developed to measure exclusive breastfeeding self-efficacy and is available for Acholi and Langi-speaking women (Boateng et al., 2019).

The PBSES, developed to measure women's perception of breastfeeding self-efficacy before delivery, has been validated in three languages (Pineiro-Albero et al., 2013; Aydın & Pasinlioğlu, 2018; Hazar & Akça, 2018; Wells et al., 2006). The short form of PBSES (PBSES-SF) has been developed in Spanish (Silva-Tubio et al., 2021). Recently, McKinley et al. (2019) developed the PREPBS, a comprehensive self-report tool for measuring a woman's perceived self-efficacy to engage in breastfeeding behaviors after giving birth. Finally, the BPEBI was developed as a self-report tool to measure breastfeeding personal efficacy beliefs at any time before and after giving birth; the BPEBI has been tested in multi-ethnic volunteer female students in the US to sustain health promotion interventions (Cleveland & Susan McCrone, 2005).

**Table 2** displays the measurement properties of each tool. The BSES revealed a two-factor solution with 31-34 items on a five-point Likert scale, with varying assessment time-points from one week to four months postpartum (Creedy et al., 2003; Eksioglu & Ceber, 2011). The BSES-SF is a mono-dimensional tool with 14 items, tested antenatally and up to twelve-month postpartum (Brandão et al., 2018; Dennis et al., 2011; Iliadou et al., 2020; Tokat et al., 2010; Basu et al., 2020). The BSES-EBF has a two-factor solution, with nine fewer items than BSES-SF, tested up to three months postpartum (Boateng et al., 2019). The PBSES exhibits a mono-dimensional second-order solution with four latent dimensions (Hazar & Akça, 2018; Pineiro-Albero et al., 2013), while two studies described a single-factor solution (Aydin et al., 2017; Wells et al., 2006). The tool consists of 19-20 items on a 5-point Likert scale (Aydın & Pasinlioğlu, 2018; Hazar & Akça, 2018; Pineiro-Albero et al., 2013; Wells et al., 2006). The PBSES-SF has a second-order factor with three latent dimensions and twelve items (Silva-Tubio et al., 2021). The PREPBS features a four-factor solution with 39 items on a 10-point response scale (Mckinley et al., 2019). Lastly, the BPEBI comprises a five-factor solution and 27 visual analog scale items ranging from 0% to 100%.

## Quality methodological assessment

The quality methodological assessment is reported in **Table 3**.

### *Content validity*

Most studies had doubtful methodological quality due to unclear descriptions of relevance, comprehensiveness, and comprehensibility assessment methods (e.g., Asgarian et al., 2019; Aydın & Pasinlioğlu, 2018; Basu et al., 2020). Relevance was assessed by asking patients in six studies (e.g., Dai & Dennis, 2003; Eksioglu & Ceber, 2011) and professionals in 14 studies (e.g., Aydın & Pasinlioğlu, 2018; Boateng et al., 2019). Patient comprehensibility was evaluated in 18 studies (e.g., Brandão et al., 2018; Cleveland & Susan McCrone, 2005), professional comprehensiveness in six (e.g., Cleveland & Susan McCrone, 2005; Dai & Dennis, 2003), and patient comprehensiveness in seven studies (e.g., Dai & Dennis, 2003; Dennis & Faux, 1999).

Sufficient content validity evidence (+) was provided in 18 studies (e.g., Aydın & Pasinlioğlu, 2018; Bosnjak et al., 2012; Brandão et al., 2018), but not tested in 11 studies (e.g., Amini et al., 2019; Creedy et al., 2003; Dennis, 2003). Two studies with adequate or very good methodological quality scores still had insufficient (-) or indeterminate (?) evidence quality (Hazar & Akça, 2018; Radwan et al., 2022).

### *Structural validity*

All studies assessed structural validity, but only Nanishi et al. (2015) provided no evidence. In 29 studies, methodological quality scores were adequate or very good (e.g., Amini et al., 2019; Aydın & Pasinlioğlu, 2018; Basu et al., 2020), with appropriate factor analysis and sample sizes. Six studies had doubtful methodological quality scores (e.g., Gregory et al., 2008; Oriá et al., 2009; Pavicic Bosnjak et al., 2012). Regarding evidence quality, eight studies were rated insufficient (-) due to explained variance lower than 50% (e.g., Asgarian et al., 2019; Aydın & Pasinlioğlu, 2018; Dai & Dennis, 2003), and ten studies were rated indeterminate (?) when explained variance was not mentioned (e.g., Amini et al., 2019; Gregory et al., 2008; McCarter-Spaulding & Dennis, 2010).

### *Internal Consistency*

Per COSMIN guidelines, Cronbach's alpha is recommended to test reliability. In eight studies, methodological quality scores were inadequate as Cronbach's alpha was provided only for overall dimensions (e.g., Basu et al., 2020; Brandão et al., 2018; Cleveland & Susan McCrone, 2005). However, 28 studies had sufficient evidence quality (+), reporting Cronbach's alpha ≥ 0.70 for each sub-dimension (e.g., Amini et al., 2019; Asgarian et al., 2019; Aydın & Pasinlioğlu, 2018; Boateng et al., 2019; Dennis, 2003).

***Cross-cultural validity***

The overall methodological quality score in fourteen studies was adequate, using suitable analysis approaches and involving similar target groups for the main characteristics (e.g., Amini et al., 2019; Asgarian et al., 2019). However, the quality of evidence was rated indeterminate (?) for all studies because the cross-cultural validity was not tested using the measurement invariance approach.

***Reliability***

Test-retest reliability was assessed in ten studies, with adequate methodological quality in five (e.g., Boateng et al., 2019; Creedy et al., 2003). Quality was doubtful in studies lacking justification for interval time, evidence of sample stability, or adequate reliability testing (e.g., Aydın & Pasinlioğlu, 2018; Economou et al., 2021). Sufficient evidence quality (+) was found in three studies with intraclass correlation (ICC) ≥ 0.70 or Pearson correlation ≥ 0.80 (e.g., Aydın & Pasinlioğlu, 2018; Mckinley et al., 2019).

***Hypothesis testing for construct validity***

The overall methodological quality score was adequate or very good for most studies, but six lacked evidence on hypothesis testing (e.g., Aydın & Pasinlioğlu, 2018; Eksioglu & Ceber, 2011). Known-groups validity in twenty-one studies showed that women with prior, positive breastfeeding experiences and willingness to breastfeed had higher self-efficacy levels than first-time mothers with low literacy and social support. Some studies reported positive and significant convergent validity with similar tools (e.g., BSES with QMIDAT and H & H Lactation Scale).

Negative and significant discriminant validity was found with different tools (e.g., BSES-SF with EPDS and STAI). Insufficient evidence quality was only rated for studies where hypothesis testing was unconfirmed (e.g., Asgarian et al., 2019; Zubaran et al., 2016).

***Criterion validity***

In this case, a sub-form of predictive criterion validity was tested using breastfeeding behaviors during postpartum as the gold standard. Nine studies had adequate to very good overall methodological quality scores for correlational analysis or sensitivity analysis (e.g., Boateng et al., 2019; Economou et al., 2021). For BSES-SF and PBSES-SF, cutoff points for maximum performance predicting breastfeeding status ranged between 45-50 and 40 points, respectively. However, 20 studies had inadequate or doubtful overall methodological quality scores (e.g., Basu et al., 2020; Brandão et al., 2018), and seven studies did not provide criterion validity (e.g., Amini et al., 2019; Asgarian et al., 2019). The quality of evidence was sufficient for seven studies that showed significant correlation with the gold standard or predictive power detection of the score ≥ 70% (e.g., Boateng et al., 2019; Economou et al., 2021; Nanishi et al., 2015).

***Rating and grading***

The overall qualitative rating was performed for breastfeeding self-efficacy tools with more than two studies, and the results are reported in the Summary of Findings (**Table 4**). Four tools were not graded due to insufficient evidence (PBSES-SF; PREPBS; BSES-EBF; BPEBI).

**BSES**. This tool had moderate quality evidence for content validity, structural validity, and hypothesis testing for construct validity. However, it had low-quality evidence for internal consistency and inconsistent ratings for cross-cultural validity, reliability, and criterion validity. According to COSMIN guidelines, BSES is recommended for measuring a mother's confidence in her breastfeeding ability after birth (Grade A).

**BSES-SF**. This tool had moderate quality evidence for content validity, structural validity, and hypothesis testing for construct validity. It had high-quality evidence for internal consistency and

criteria validity. However, it had inconsistent ratings for cross-cultural validity and reliability. According to COSMIN guidelines, BSES-SF is recommended for use (Grade A).

**PBSES**. This tool had moderate quality evidence for content validity, structural validity, reliability, and hypothesis testing for construct validity. It had high-quality evidence for internal consistency. However, it had inconsistent ratings for cross-cultural validity and criterion validity. According to COSMIN guidelines, PBSES is recommended for measuring women's perception of breastfeeding self-efficacy before delivery (Grade A).

## Discussion

This systematic review aimed to assess and summarize the quality of evidence regarding the measurement of breastfeeding self-efficacy tools using the COSMIN guideline (Mokkink et al., 2018). Although a critical review of the available tools on breastfeeding self-efficacy was preliminarily performed by Tuthill et a. (2016), a descriptive approach was used to summarize the available evidence (Tuthill et al., 2016). Therefore, adopting a COSMIN methodology approach enables authors to improve the selection of the most suitable tools and provide up-to-date evidence. The authors agree to "recommend" the BSES, BSES-SF, and PBSES (Grade A) as valid and reliable self-report tools to measure breastfeeding self-efficacy in the continuum of maternity care.

Among the recommended tools of breastfeeding self-efficacy (Grade A), the BSES-SF is the most feasible tool to use in clinical practice and the most utilized internationally. Therefore, the BSES-SF allows women to self-assess their self-efficacy with a reduced burden related to the need to answer several items, acknowledging that it is a short form and is available in fifteen languages. Accordingly, the response rate achieved for the BSES-SF ranged from 54% to 86% in the postpartum follow-up (Dennis, 2003; Dennis et al., 2011; Economou et al., 2021; Gerhardsson et al., 2014; Gregory et al., 2008). For the same reasons, the PBSES-SF could facilitate the measurement of breastfeeding self-efficacy before delivery. However, for the PBSES-SF and each tool based on one validation study, it was not feasible to provide grading of evidence using the GRADE-modified approach (Schünemann et al., 2013). This was the reason underpinning the

choice to perform only a methodological quality evaluation of the measurement properties for the following tools: PBSES-SF, PREPBS, BSES-EBF, and BPEBI.

The BSES-EBF tool was recently developed by Boateng et al. (2019) as an adaptation of the BSES-SF to assess the mother's confidence to practice exclusive breastfeeding for up to six months (Boateng et al., 2019). Furthermore, the fact that the BSES-EBF encompassed a few items (9 items) enhances its usability for clinical contexts to identify women at more risk of inadequate breastfeeding behaviours. However, when it is needed to address educational initiatives on breastfeeding initiation, continuation, and perseverance, the BPEBI is indicated because it measures women's personal beliefs about the breastfeeding experience. In this scenario, further studies are needed to increase the validity of these tools, particularly in those cultural settings where exclusive breastfeeding is far from reaching the standard goal defined by WHO (WHO/UNICEF, 2019).

Specifically for the measurement properties of the "recommended" tools (i.e., BSES, BSES-SF, PBSES), moderate quality of evidence for sufficient content and construct validity (i.e., structure and hypotheses testing for construct validity property) was defined for all the analyzed tools. Content and construct validity are the most pivotal measurement properties due to testing the relevance, comprehensiveness, and comprehensibility of the self-report tools considering the users' perspectives, as well as defining the psychometric structure of the tools (Magon et al., 2023; Mikkonen et al., 2022).

In most of the studies, multiparous women with prior breastfeeding experience, high education level, and adequate social support showed higher levels of self-efficacy than the primiparous (Aydın & Pasinlioğlu, 2018; Eksioglu & Ceber, 2011; Hazar & Akça, 2018; Iliadou M. et al., 2020; Nanishi et al., 2015; Wan-Yim et al., 2012). Therefore, we are confident that the validity of these tools could ensure distinguishing women with higher breastfeeding self-efficacy levels based on specific social-demographic characteristics and previous breastfeeding experience. In this scenario, future educational interventions should be informed by BSES theory to deliver tailored interventions.

The literature acknowledges the predictive role of self-efficacy in enhancing breastfeeding rates. High-quality evidence for sufficient criterion validity was identified only for the BSES-SF. More specifically, the gold-standard criteria used to test the predictive validity of the BSES-SF was the status of breastfeeding (defined as exclusively breastfeeding, partially breastfeeding, or using bottle-feeding) in the postpartum (Labbok & Krasovec, 1990): lower levels of breastfeeding self-efficacy in the first months negatively predict a practice of exclusive breastfeeding over time. Therefore, breastfeeding challenges mainly occur in the first month, and that can be related to social factors (e.g., low social support or the return to work) or clinical factors (postpartum depression, insufficient milk supply, fatigue, pain, or cracked nipples) (Gianni et al., 2019). In this scenario, the BSES-SF is the only tool useful to identify mothers at high risk of suspending breastfeeding prematurely.

The results are inconsistent for the reliability property for the BSES and BSES-SF and indeterminate for the cross-cultural validity property for all the included tools. Thus, the authors acknowledged low confidence in stating the adequate stability of the tools under different conditions (i.e., antenatal and postpartum) and cultural settings. Therefore, the measurement invariance of the tools over time and among different cultural groups of patients has been thus far under-investigated. These results are consistent with the critical review of Tuthill et al. (2016), highlighting that future studies should address the accuracy of the tools in measuring self-efficacy breastfeeding longitudinally (Tuthill et al., 2016). In this scenario, more evidence regarding the responsiveness of the tools is needed to ensure their efficacy in detecting changes in self-efficacy breastfeeding levels over time and in identifying the effectiveness of the educational intervention to increase the mothers' confidence to breastfeed (Mokkink et al., 2018).

**Limitations**

This study has several limitations. Firstly, a potential limitation of this systematic review is that we only searched four databases, which may have limited our ability to retrieve all relevant literature on the topic. This employed approach could potentially introduce a publication bias, as

studies that were not included in these databases may have different results. Additionally, we restricted our search to papers published in English or Italian, which may have further limited our ability to capture all relevant literature. As a result, the external validity of our findings may be limited. However, we aimed to minimize the influence of these limitations by conducting a comprehensive search within the area defined by our inclusion and exclusion criteria. In this study, we have yet to consider the quality of the developmental phase of each tool because it is not a measurement property. However, according to Bandura's social cognitive theory, all the studies provided a clear and common description of the construct to be measured. In addition, a lack of consistency in reporting the results from the included studies made it challenging to compare and evaluate different breastfeeding self-efficacy tools across studies because studies may not have reported on all of the necessary measurement properties, or they may not have reported the information in a consistent or clear manner. As a result, it was not possible to fully evaluate the measurement properties of the PROM using the COSMIN checklist. Finally, only a qualitative synthesis of evidence for the other measurement proprieties was feasible in this systematic review because the heterogeneity of the factor structure did not allow researchers to pool quantitative measures.

## Conclusion

This systematic review allowed summarizing the psychometric properties of the available self-report breastfeeding self-efficacy tools. Thus, the results of this study could be helpful for clinicians and researchers. As per clinicians, using in routine clinical practice breastfeeding self-efficacy tools with high certainty of evidence (i.e., Grade A) is recommended to measure and support breastfeeding self-efficacy practice correctly. Research should address some of the weakest measurement properties identified in this study. Overall, the measurement of breastfeeding self-efficacy and the resulting educational interventions are recommended to increase the rate of exclusive breastfeeding.

# References

*Amini, P., Samani, R. O., Sepidarkish, M., Hashiani, A. A., Hosseini, M., & Maroufizadeh, S. (2019). The Breastfeeding Self-Efficacy Scale-Short Form (BSES-SF): A validation study in Iranian mothers. *BMC Research Notes*, *12*(1). https://doi.org/10.1186/s13104-019-4656-7

*Asgarian, A., Hashemi, M., Pournikoo, M., Mirazimi, T., Zamanian, H., & Mohammadali, A.-T. (2019). Translation , Validation , and Psychometric Properties of Breastfeeding Self-Efficacy Scale—Short Form Among Iranian Women. *Journal of Human Lactation*, *36*(2), 227–235. https://doi.org/10.1177/0890334419883572

*Aydın, A., & Pasinlioğlu, T. (2018). Reliability and Validity of a Turkish version of the Prenatal Breastfeeding Self-Efficacy Scale. *Midwifery*, *64*, 11–16. https://doi.org/10.1016/j.midw.2018.05.007

Bandura, A. (1977). Self-Efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, *84*, 191–215.

Bandura, A. (2001). Social cognitive theory: An agentic perspective. *Annual Review of Psychology*, *52*(1), 1–26. https://doi.org/10.1146/annurev.psych.52.1.1

Bandura, Freeman, W. H., & Lightsey, R. (1999). Self-Efficacy: The Exercise of Control. *Journal of Cognitive Psychotherapy*, *13*(2), 158–166. https://doi.org/10.1891/0889-8391.13.2.158

*Basu, S., Garg, S., Sharma, A., Arora, E., & Singh, Mm. (2020). The hindi version of the breastfeeding self-efficacy scale-short form: Reliability and validity assessment. *Indian Journal of Community Medicine*, *45*(3), 348. https://doi.org/10.4103/ijcm.ijcm_378_19

*Boateng, G. O., Martin, S. L., Tuthill, E. L., Collins, S. M., Dennis, C.-L., Natamba, B. K. ù, & Young, S. L. (2019). Adaptation and psychometric evaluation of the breastfeeding self-efficacy scale to assess exclusive breastfeeding. *BMC pregnancy and childbirth*, *19*(73).

*Bosnjak, A. P., Rumboldt, M., Stanojevic, M., & Dennis, C.-L. (2012). Psychometric Assessment of the Croatian Version of the Breastfeeding Self-Efficacy Scale – Short Form. *Journal of Human Lactation*, *28*(4), 565–569. https://doi.org/10.1177/0890334412456240

*Brandão, A. B. P., De Abreu, I. C., Aimbire, F., Higa, E. M., Casali, A., Ferreira, F. G., Albuquerque, R. C. M., Santos, L. B., Irigoyen, M. C. C., Casali, K. R., & Cunha, T. S. (2018). Saccharomyces boulardii attenuates autonomic cardiovascular dysfunction and modulates inflammatory cytokines in diabetic mice. *Diabetes*, *67*((Brandão A.B.P.; De Abreu I.C.; Aimbire F.; Higa E.M.; Casali A.; Ferreira F.G.; Albuquerque R.C.M.; Santos L.B.; Irigoyen M.C.C.; Casali K.R.; Cunha T.S.) São José dos Campos, Brazil, São Paulo, Brazil), A617. Embase.

*Cleveland, A. P., & McCrone, S. (2005). Development of the Breastfeeding Personal Efficacy Beliefs Inventory: A Measure of Women's Confidence About Breastfeeding. *Journal of Nursing Measurement*, *13*(2), 115–127.

*Creedy, D. K., Dennis, C.-L., Blyth, R., Moyle, W., Pratt, J., & Vries, S. M. D. (2003). Psychometric Characteristics of the Breastfeeding Self-Efficacy Scale: Data from an Australian Sample. *Research in Nursing & Health*, *26*, 143–152. https://doi.org/10.1002/nur.10073

*Dai, X., & Dennis, C.-L. (2003). Translation and Validation of the Breastfeeding Self-Efficacy Scale Into Chinese. *Journal of Midwifery and Women's Health*, *48*(5), 350–356. https://doi.org/10.1053/S1526-9523(03)00283-6

De Roza, J. G., Fong, M. K., Ang, B. L., Sadon, R. B., Koh, E. Y. L., & Teo, S. S. H. (2019). Exclusive breastfeeding, breastfeeding self-efficacy and perception of milk supply among mothers in Singapore: A longitudinal study. *Midwifery*, *79*, 102532. https://doi.org/10.1016/j.midw.2019.102532

*Dennis, C.-L. (1999). Theoretical Underpinnings of Breastfeeding Confidence: A Self-Efficacy Framework. *Journal of Human Lactation*, *15*(3), 195–201.

*Dennis, C.-L. (2003). The Breastfeeding Self-Efficacy Scale: Psychometric Assessment of the Short Form. *Journal of Obstetric, Gynecologic & Neonatal Nursing*, *32*(6), 734–744.

Dennis, C.-L., Brennenstuhl, S., & Dick, J. A. (2018). Measuring Paternal Breastfeeding Self-Efficacy: A Psychometric Evaluation of the Breastfeeding Self-Efficacy Scale—Short Form among Fathers. *Midwifery*, *64*, 17–22.

Dennis, C.-L., & Faux, S. (1999). Development and psychometric testing of the Breastfeeding Self-Efficacy Scale. *Research in Nursing & Health Health*, *22*(5), 399–409. https://doi.org/10.1002/(SICI)1098-240X(199910)22:5<399::AID-NUR6>3.0.CO;2-4

*Dennis, C.-L., Heaman, M., & Mossman, M. (2011). Psychometric Testing of the Breastfeeding Self-Efficacy Scale-Short Form Among Adolescents. *Journal of Adolescent Health*, *49*(3), 265–271. https://doi.org/10.1016/j.jadohealth.2010.12.015

*Economou, M., Kolokotroni, O., Paphiti-Demetriou, I., Kouta, C., Lambrinou, E., Hadjigeorgiou, E., Hadjiona, V., & Middleton, N. (2021). The association of breastfeeding self-efficacy with breastfeeding duration and exclusivity: Longitudinal assessment of the predictive validity of the Greek version of the BSES-SF tool. *BMC Pregnancy and Childbirth*, *21*(1). https://doi.org/10.1186/s12884-021-03878-3

*Eksioglu, A. B., & Ceber, E. (2011). Translation and validation of the Breast-feeding Self-efficacy Scale into Turkish. *Midwifery*, *27*, e246–e253. https://doi.org/10.1016/j.midw.2010.10.009

Elsman, E. B. M., Butcher, N. J., Mokkink, L. B., Terwee, C. B., Tricco, A., Gagnier, J. J., Aiyegbusi, O. L., Barnett, C., Smith, M., Moher, D., & Offringa, M. (2022). Study protocol for developing, piloting and disseminating the PRISMA-COSMIN guideline: A new reporting guideline for systematic reviews of outcome measurement instruments. *Systematic Reviews*, *11*(1), 121. https://doi.org/10.1186/s13643-022-01994-5

Gavine, A., Shinwell, S. C., Buchanan, P., Farre, A., Wade, A., Lynn, F., Marshall, J., Cumming, S. E., Dare, S., & McFadden, A. (2022). Support for healthy breastfeeding mothers with healthy term babies. *The Cochrane Database of Systematic Reviews*, *10*(10), CD001141. https://doi.org/10.1002/14651858.CD001141.pub6

Gerçek, E., Karabudak, S. S., Çelik, N. A., & Saruhan, A. (2017). The relationship between breastfeeding self-efficacy and LATCH scores and affecting factors. *Journal of Clinical Nursing*, *26*(7–8), 994–1004. https://doi.org/10.1111/jocn.13423

*Gerhardsson, E., Nyqvist, K. H., Mattsson, E., Volgsten, H., Hildingsson, I., & Funkquist, E. (2014). The

Swedish Version of the Breastfeeding Self-Efficacy Scale – Short Form: Reliability and Validity Assessment. *Journal of Human Lactation*, *30*(3), 340–345. https://doi.org/10.1177/0890334414523836

Gianni, M. L., Bettinelli, M. E., Manfra, P., Sorrentino, G., Bezze, E., Plevani, L., Cavallaro, G., Raffaeli, G., Crippa, B. L., Colombo, L., Morniroli, D., Liotto, N., Roggero, P., Villamor, E., Marchisio, P., & Mosca, F. (2019). Breastfeeding difficulties and risk for early breastfeeding cessation. *Nutrients*, *11*(10). https://doi.org/10.3390/nu11102266

*Gregory, A., Penrose, K., Morrison, C., Dennis, C.-L., & MacArthur, C. (2008). Psychometric Properties of the Breastfeeding Self-Efficacy Scale- Short Form in an Ethnically Diverse U.K. Sample. *Public Health Nursing*, *25*(3), 278–284. https://doi.org/10.1111/j.1525-1446.2008.00705.x

*Hazar, H. U., & Akça, E. U. (2018). Prenatal breastfeeding self efficacy scale: Validity and reliability study. *Turk Pediatri Ars*, *53*(4), 222–230. https://doi.org/10.5152/TurkPediatriArs.2018.18114

*Iliadou, M., Lykeridou, K., Prezerkos, P., Zyga, S., Sakellari, E., Vivilaki, V., & Tziaferi, S. (2020). Psychometric properties of the Greek version of the Breastfeeding Self-Efficacy Scale and correlation with depressive symptomatology. *Archives of Hellenic Medicine*, *37*(5), 656–662.

Ingram, J., Johnson, D., Copeland, M., Churchill, C., & Taylor, H. (2015). The development of a new breast feeding assessment tool and the relationship with breast feeding self-efficacy. *Midwifery*, *31*(1), 132–137. https://doi.org/10.1016/j.midw.2014.07.001

Kronborg, H., & Væth, M. (2019). Validation of the breastfeeding score—A simple screening tool to predict breastfeeding duration. *Nutrients*, *11*(12). https://doi.org/10.3390/nu11122852

Labbok, M., & Krasovec, K. (1990). Toward consistency in breastfeeding definitions. *Studies in Family Planning*, *21*(4), 226–230.

Magon, A., Conte, G., Dellafiore, F., Arrigoni, C., Baroni, I., Brera, A. S., Avenido, J., De Maria, M., Stievano, A., Villa, G., & Caruso, R. (2023). Nursing Profession Self-Efficacy Scale—Version 2: A Stepwise Validation with Three Cross-Sectional Data Collections. *Healthcare*, *11*(5), Art. 5. https://doi.org/10.3390/healthcare11050754

*McCarter-Spaulding, D. E., & Dennis, C.-L. (2010). Psychometric Testing of the Breastfeeding Self-Efficacy Scale-Short Form in a Sample of Black Women in the United States. *Research in Nursing & Health*, *33*, 111–119. https://doi.org/10.1002/nur.20368

*Mckinley, E. M., Knol, L. L., Turner, L. W., Burnham, J. J., Graettinger, K. R., Hernandez-reif, M., & Leeper, J. D. (2019). The Prenatal Rating of Efficacy in Preparation to Breastfeed Scale: A New Measurement Instrument for Prenatal Breastfeeding Self-Efficacy. *Journal of Human Lactation*, *35*(1), 21–31. https://doi.org/10.1177/0890334418799047

*McQueen, K. A., Montelpare, W., & Dennis, C.-L. (2013). Breastfeeding and Aboriginal Women: Validation of the Breastfeeding Self- Efficacy Scale – Short Form. *Canadian Journal of Nursing Research*, *45*(2). https://doi.org/10.1177/084456211304500209

Mikkonen, K., Tomietto, M., & Watson, R. (2022). Instrument development and psychometric testing in nursing education research. *Nurse Education Today*, *119*, 105603. https://doi.org/10.1016/j.nedt.2022.105603

Mokkink, L. B., de Vet, H. C. W., Prinsen, C. a. C., Patrick, D. L., Alonso, J., Bouter, L. M., & Terwee, C. B. (2018). COSMIN Risk of Bias checklist for systematic reviews of Patient-Reported Outcome Measures. *Quality of Life Research: An International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation*, *27*(5), 1171–1179. https://doi.org/10.1007/s11136-017-1765-4

Mokkink, L. B., Prinsen, C. A., Patrick, D. L., Alonso, J., Bouter, L. M., Vet, H. C. de, & Terwee, C. B. (2018). *COSMIN methodology for systematic reviews of Patient-Reported Outcome Measures (PROMs). User manual*.

Mokkink, L. B., Terwee, C. B., Knol, D. L., Stratford, P. W., Alonso, J., Patrick, D. L., Bouter, L. M., & de Vet, H. C. W. (2006). Protocol of the COSMIN study: COnsensus-based Standards for the selection of health Measurement INstruments. *BMC Medical Research Methodology*, *6*, 2. https://doi.org/10.1186/1471-2288-6-2

Mokkink, L. B., Terwee, C. B., Prinsen, C. A. C., de Vet, H. C. W., & COSMIN Steering Committee.

(2016). Taxonomy of measurement properties: A commentary on Polit (2015). *International Journal of Nursing Studies*, *53*, 399–400. https://doi.org/10.1016/j.ijnurstu.2015.08.010

*Nanishi, K., Green, J., Taguri, M., & Jimba, M. (2015). Determining a Cut-Off Point for Scores of the Breastfeeding Self-Efficacy Scale – Short Form: Secondary Data Analysis of an Intervention Study in Japan. *PLOS ONE*, *10*(6), 1–12. https://doi.org/10.1371/journal.pone.0129698

*Oliver-Roig, A., D'Anglade-Gonzalez, M.-L., Garcia-Garcia, B., Silva-Tubio, J.-R., Richart-Martınez, M., & Dennis, C.-L. (2012). The Spanish version of the Breastfeeding Self-Efficacy Scale-Short Form: Reliability and validity assessment. *International Journal of Nursing Studies*, *49*, 169–173.

*Oriá, O. B. M., Ximenes, L. B., Almeida, P. C. de, Glick, D. F., & Dennis, C.-L. (2009). Psychometric Assessment of the Brazilian Version of the Breastfeeding Self-Efficacy Scale. *Public Health Nursing*, *26*(6), 574–583. https://doi.org/10.1111/j.1525-1446.2009.00817.x

Palmér, L., & Jutengren, G. (2019). Development and psychometric testing of an instrument to assess existential aspects of mother's initial breastfeeding difficulties (ExBreastS). *Sexual and Reproductive Healthcare*, *19*, 88–94. https://doi.org/10.1016/j.srhc.2019.01.005

*Petrozzi, A., & Gagliardi, L. (2016). Breastfeeding Self-Efficacy Scale: Validation of the Italian Version and Correlation With Breast-feeding at 3 Months. *Journal of Pediatric Gastroenterology and Nutrition*, *62*(1), 137–139. https://doi.org/10.1097/MPG.0000000000000902

*Pineiro-Albero, R. M., Ramos-Pichardo, J. D., Oliver-Roig, A., Miguel Velandrino-Nicolas, A. R.-M., Garcıa-de-Leon-Gonzalez, R., & Wells, K. J. (2013). The Spanish version of the Prenatal Breast-feeding Self-efficacy Scale: Reliability and validity assessment. *International Journal of Nursing Studies*, *50*, 1385–1390. https://doi.org/10.1016/j.ijnurstu.2012.12.010

Prinsen, C. A., Vohra, S., Rose, M. R., Boers, M., Tugwell, P., Clarke, M., Williamson, P. R., & Terwee, C. B. (2016). How to select outcome measurement instruments for outcomes included in a "Core Outcome Set"—A practical guideline. *Trials*, 17(1), 449. https://doi.org/10.1186/s13063-016-1555-2

Prinsen, C. a. C., Mokkink, L. B., Bouter, L. M., Alonso, J., Patrick, D. L., de Vet, H. C. W., & Terwee, C.

B. (2018). COSMIN guideline for systematic reviews of patient-reported outcome measures. *Quality of Life Research: An International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation, 27*(5), 1147–1157. https://doi.org/10.1007/s11136-018-1798-3

*Radwan, H., Fakhry, R., Boateng, G. O., Metheny, N., issa, W. B., Faris, M. A. I. E., Obaid, R. S., Marzooqi, S. A., Ghazal, H. A., & Dennis, C. L. (2022). Translation and Psychometric Evaluation of the Arabic Version of the Breastfeeding Self-Efficacy Scale-Short Form Among Women in the United Arab Emirates. *Journal of Human Lactation*, 39(1), 40-50. https://doi.org/10.1177/08903344221084623

Scholz, U., Doña, B. G., Sud, S., & Schwarzer, R. (2002). Is general self-efficacy a universal construct? Psychometric findings from 25 countries. *European Journal of Psychological Assessment, 18*(3), 242–251.

Schünemann, H., Brożek, J., Guyatt, G., & Oxman, A. (2013). *GRADE handbook*. https://gdt.gradepro.org/app/handbook/handbook.html

*Silva-Tubio, J. R., Oliver-Roig, A., Perpiñá-Galvañ, J., & Richart-Martínez, M. (2021). Reliability and validity of the reduced Spanish version of the Prenatal Breastfeeding Self-Efficacy Scale. *Research in Nursing and Health, 44*(6), 979–991. https://doi.org/10.1002/nur.22179

Terwee, C. B., Prinsen, C. a. C., Chiarotto, A., Westerman, M. J., Patrick, D. L., Alonso, J., Bouter, L. M., de Vet, H. C. W., & Mokkink, L. B. (2018). COSMIN methodology for evaluating the content validity of patient-reported outcome measures: A Delphi study. *Quality of Life Research: An International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation*, *27*(5), 1159–1170. https://doi.org/10.1007/s11136-018-1829-0

*Tokat, M. A., Okumus, H., & Dennis, C.-L. (2010). Translation and psychometric assessment of the Breast-feeding Self-Efficacy Scale—Short Form among pregnant and postnatal women in Turkey. *Midwifery, 26*, 101–108. https://doi.org/10.1016/j.midw.2008.04.002

*Torres, M. M., Torres, R. R. ; D., Rodríguez, A. M. ; P., & Dennis, C.-L. (2003). Translation and Validation of the Breastfeeding Self-Efficacy Scale Into Spanish: Data From a Puerto Rican Population. *Journal of Human Lactation, 19*(1), 35–42. https://doi.org/10.1177/0890334402239732

Tuthill, E. L., McGrath, J. M., Graber, M., Cusson, R. M., & Young, S. L. (2016). Breastfeeding Self-efficacy: A Critical Review of Available Instruments. *Journal of Human Lactation*, *32*(1), 35–45. https://doi.org/10.1177/0890334415599533

UNICEF. (2018). *Breastfeeding. A Mother's Gift, for Every Child*.

*Wan-Yim, I., Lai-Shan, Y., Kai-Chow, C., Sek-Ying, C., & Cindy-Lee;, D. (2012). Translation and Validation of the Hong Kong Chinese Version of the Breastfeeding Self-Efficacy Scale—Short Form. *Research in Nursing & Health*, *35*(5), 450–459. https://doi.org/10.1002/nur.21493

*Wan-Yim, I., Ling-Ling, G., Kai-Chow, C., Janita Pak-Chun, C., & Yang, X. (2016). The Short Form of the Breastfeeding Self-Efficacy Scale as a Prognostic Factor of Exclusive Breastfeeding among Mandarin-Speaking Chinese Mothers. *Journal of Human Lactation*, *32*(4), 711–720. https://doi.org/10.1177/0890334416658014

*Wells, K. J., Thompson, N. J., & Kloeblen-Tarver, A. S. (2006). Development and pychometric testing of the prenatal breast-feeding self-efficacy scale. *American Journal of Health Behavior*, *30*(2), 177–187. https://doi.org/10.5993/AJHB.30.2.7

Wheeler, B. J., & Dennis, C.-L. (2013). Psychometric Testing of the Modified Breastfeeding Self-Efficacy Scale (Short Form) Among Mothers of Ill or Preterm Infants. *JOGNN - Journal of Obstetric, Gynecologic, and Neonatal Nursing*, *42*(1), 70–80. https://doi.org/10.1111/j.1552-6909.2012.01431.x

WHO. (2021). *Infant and young child feeding*. https://www.who.int/news-room/fact-sheets/detail/infant-and-young-child-feeding

WHO/UNICEF. (2019). *Discussion paper. The extension of the 2025 Maternal, Infant and Young Child nutrition targets to 2030*. http://www.who.int/nutrition/healthygrowthproj/en/index1.html

Williamson, P. R., Altman, D. G., Bagley, H., Barnes, K. L., Blazeby, J. M., Brookes, S. T., Clarke, M., Gargon, E., Gorst, S., Harman, N., Kirkham, J. J., McNair, A., Prinsen, C. A. C., Schmitt, J., Terwee, C. B., & Young, B. (2017). The COMET Handbook: Version 1.0. *Trials*, *18*(3), 280. https://doi.org/10.1186/s13063-017-1978-4

*Wutke, K., & Dennis, C.-L. (2007). The reliability and validity of the Polish version of the Breastfeeding Self-Efficacy Scale-Short Form: Translation and psychometric assessment. *International Journal of Nursing Studies*, *44*, 1439–1446. https://doi.org/10.1016/j.ijnurstu.2006.08.001

*Zubaran, C., Foresti, K., Schumacher, M., Thorell, M. R., Amoretti, A., Muller, L., & Dennis, C.-L. (2016). The Portuguese Version of the Breastfeeding Self-Efficacy Scale – Short Form. *Journal of Human Lactation*, *26*(3), 297–303. https://doi.org/10.1177/0890334409359916