

City Research Online

City St George's, University of London

Citation: Jarašiūnaitė-Fedosejeva, G., Radoš, S. N., Ayers, S., Brekalo, M., Costa, R., Horsch, A., Ionio, C., Kalfon-Hakhmigari, M., Pinto, T. M., Rattaz, V., et al (2026). The short version of the Postpartum Bonding Questionnaire: A cross-country psychometric study of mothers. Midwifery, 163, 105020. doi: 10.1016/j.midw.2026.105020

This is the published version of the paper.

This version of the publication may differ from the final published version. To cite this item please consult the publisher's version.

Permanent repository link: https://openaccess.city.ac.uk/id/eprint/38601/

Link to published version: https://doi.org/10.1016/j.midw.2026.105020

Copyright and Reuse: Copyright and Moral Rights remain with the author(s) and/or copyright holders. Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge, unless otherwise indicated, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way. For full details of reuse please refer to City Research Online policy.

City Research Online: http://openaccess.city.ac.uk/ publications@citystgeorges.ac.uk

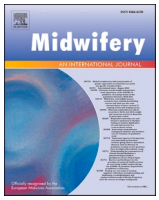
Midwifery 163 (2026) 105020



Contents lists available at ScienceDirect

Midwifery

journal homepage: www.elsevier.com/locate/midw



The short version of the Postpartum Bonding Questionnaire: A cross-country psychometric study of mothers

Gabija Jarašiūnaitė-Fedosejeva [a,1], Sandra Nakić Radoš [b,1], Susan Ayers [c], Maja Brekalo [b], Raquel Costa [d], Antje Horsch [e,f], Chiara Ionio [g], Maor Kalfon-Hakhmigari [h], Tiago Miguel Pinto [d], Valentine Rattaz [e,i,j], Olga Riklikienė [k], Valgerður Lísa Sigurðardóttir [l,m], Emma Marie Swift [l,n], Rebecca Webb [c], Jonathan Eliahu Handelzalts [h,o,*]

[a] Department of Psychology, Vytautas Magnus University, Kaunas, Lithuania
[b] University Department of Psychology, Catholic University of Croatia, Zagreb, Croatia
[c] Centre for Maternal and Child Health Research, School of Health and Medical Sciences, City St George's, University of London, London, United Kingdom
[d] Lusófona University, HEI-Lab: Digital Human-Environment Interaction Labs, Campo Grande 376, Lisboa, 1749-024, Portugal
[e] Institute of Higher Education and Research in Healthcare, University of Lausanne, Switzerland
[f] Department Woman-mother-child, Lausanne University Hospital and University of Lausanne, Switzerland
[g] Unità di Ricerca sulla Psicologia del Trauma, Dipartimento di Psicologia, Università Cattolica, Milano, Italy
[h] School of Behavioral Sciences, The Academic College of Tel Aviv-Yafo, Tel-Aviv, 68114, Israel
[i] Center for Family Studies, Department of Psychiatry, Lausanne University Hospital and University of Lausanne, Lausanne, Switzerland
[j] Faculty of Psychology and Educational Sciences, University of Geneva, Geneva, Switzerland
[k] Department of Nursing, Faculty of Nursing, Lithuanian University of Health Sciences, Kaunas, Lithuania
[l] University of Iceland, School of Health Sciences, Faculty of Nursing and Midwifery, Reykjavík, Iceland
[m] Landspítali National University Hospital, Reykjavík, Iceland
[n] Reykjavík Birth Center, Reykjavík, Iceland
[o] Department of Psychiatry, University of Michigan, Ann Arbor, MI, USA

ARTICLE INFO

Keywords:
Postpartum bonding
Short version
Cross-cultural psychometrics
Mother
Infant

ABSTRACT

Background: The Postpartum Bonding Questionnaire (PBQ) is widely used and considered a robust measure in terms of psychometric characteristics to assess difficulties in mother–infant bonding; however, the original 25-item version has yielded inconsistent factor solutions across studies and cultural contexts. Shorter versions may enhance feasibility for routine screening. Using multicountry data, the present study aimed to establish a short version of the PBQ that reflects the original structure and demonstrates adequate psychometric properties across countries.

Methods: Data were collected from 1731 postpartum women across five European countries participating in the INTERSECT study. Participants completed the PBQ and the Edinburgh Postnatal Depression Scale at 6–12 weeks postpartum, with a subsample reassessed at 26 weeks postpartum (n = 352). Item reduction and factor structure were examined using exploratory and confirmatory factor analyses on split samples. Measurement invariance, internal consistency, test–retest reliability, and associations with depressive symptoms were assessed.

Results: Analyses supported a unidimensional 10-item PBQ version with good model fit (RMSEA= .072; CFI=.973, TLI=.965, SRMR=.052), internal consistency (McDonald's ω=.82), and test–retest reliability (ICC=.73). This version showed a strong correlation with the original PBQ and demonstrated configural, metric, and scalar measurement invariance across countries. The short version was moderately associated with depressive symptoms.

Conclusions: The findings support a brief, reliable PBQ version for assessing maternal bonding difficulties and suggest that variability in prior factor structures reflects differences in item stability rather than the absence of a coherent underlying construct.

* Corresponding author at: School of Behavioral Sciences, The Academic College of Tel Aviv-Yafo, Tel-Aviv, 68114, Israel.
E-mail address: yonatanh@mta.ac.il (J.E. Handelzalts).
[1] Shared first authorship.

https://doi.org/10.1016/j.midw.2026.105020
Received 25 June 2026; Received in revised form 14 September 2026; Accepted 18 October 2026
Available online 19 October 2026
0266-6138/© 2026 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Introduction

Statement of significance

Problem or Issue	What is Already Known	What this Paper Adds
The Postpartum Bonding Questionnaire (PBQ) is widely used to assess mother–infant bonding difficulties, but studies across countries have reported inconsistent factor structures and item compositions, limiting comparability and routine use.	Previous validation studies have identified one-, two-, three-, four-, and five-factor solutions, often requiring different item reductions. Shorter unidimensional versions have been proposed, but cross-country measurement invariance has rarely been examined.	Using data from 1731 mothers across five European countries, this study developed a brief 10-item PBQ with good psychometric properties and demonstrated configural, metric, and scalar invariance across countries, supporting its potential use in routine postpartum care, screening, and international research.

Parent-infant bonding is defined as an emotional, behavioural, cognitive, and neurobiological tie that a mother or father forms with the infant (Nakić Radoš et al., 2024). It is a parent-driven process that begins with pregnancy planning and continues through pregnancy and infancy (Bicking Kinsey and Hupcey, 2013; Nakić Radoš et al., 2024). Strong relationships with parents are vital for infant development (e.g., Brockington, 2004; Wittkowski et al., 2020). When a mother forms a positive and affectionate relationship with an infant, responding to the infant's needs with maternal sensitive behaviours, it leads to secure attachment (Deans, 2020; Montealegre-Ramón et al., 2024) and improves the infant's social and emotional development and competence, their adaptive behaviour, and cognitive outcomes (Le Bas et al., 2022). Postpartum bonding difficulties can vary in their presentation and severity, ranging from delayed or reduced positive feelings, lack of affection, and emotional distance to irritation, anger, rejection, and, in more severe cases, thoughts or behaviors related to harming the infant (Brockington, 2004; Nakić Radoš et al., 2024). Although these experiences differ in their content and severity, it remains unclear whether they represent distinct dimensions of bonding difficulties or different manifestations of a broader underlying construct. Such difficulties can lead to various negative outcomes for an infant, including diminished physical, cognitive, and socio-emotional development (e.g., Johnson, 2013).

Mother-infant bonding difficulties can be assessed using interview-based measures, such as the Stafford Interview (Brockington et al., 2017), or self-report questionnaires. A systematic review identified ten measures of parent-infant relationship (Wittkowski et al., 2020), of which the Postpartum Bonding Questionnaire (Brockington et al., 2001, 2006) was considered the most robust in terms of psychometric characteristics. The Postpartum Bonding Questionnaire (PBQ) was designed to identify mothers with early mother-infant relationship disorders, including delayed maternal emotional response toward the infant, neglect, or even abuse (Brockington et al., 2001, 2006). It consists of 25 items covering four domains: general factor, rejection and pathological anger, anxiety about the infant, and abuse (Brockington et al., 2006). The PBQ was psychometrically tested across a number of countries, with variability in its factor structure. The four-factor structure was replicated in Bangladesh (Edhborg et al., 2011), Germany (Göbel et al., 2024), Japan (Suetsugu et al., 2015), Mexico (Palacios-Hernández et al., 2026), and Spain (Garcia-Esteve et al., 2016), but often with a reduced number of items based on the analysis.

However, psychometric studies conducted in different countries have often failed to replicate the original structure, resulting in a range of alternative factor solutions. In India, a five-factor structure comprising 19 of the original 25 items was identified, consisting of the general bonding, frustration, anxiety, feeling trapped, and aggression/rejection dimensions (Vengadavaradan et al., 2019). A three-factor structure was identified in Brazil (Baldisserotto et al., 2023), Croatia (Anđelinović et al., 2019), England (Wittkowski et al., 2010), Italy (Busonera et al., 2017), and Nepal (Pudasainee-Kapri et al., 2024), as well as in a separate Japanese study (Ohashi et al., 2016). However, the content of the three factors differed across studies and included dimensions such as impaired bonding, anxiety about care and maternal distress, and lack of enjoyment and affection toward the infant. In addition, a two-factor structure was found in Israel, consisting of caregiver emotions directed to the baby and caregiver perceptions of the baby (Kalfon-Hakhmigari et al., 2025).

A number of studies have also found a one-factor solution, but with a shorter version, ranging from 8 to 22 items, such as in Brazil (Saur et al., 2023), France (Demanche et al., 2021), Japan (Kaneko and Honjo, 2014), Portugal (Nazaré et al., 2012), Türkiye (Dişsiz et al., 2024), and the USA (Bicking-Kinsey et al., 2014; Lavallée et al., 2026). Taken together, these findings suggest that shorter unidimensional versions of the PBQ may assess a general dimension of mother–infant bonding difficulties, rather than bonding disorders proposed in the original conceptualization of the PBQ.

Some of these inconsistencies in factor structure and number of items may be attributable to different sampling strategies. While most studies included community samples, several were conducted in clinical populations (e.g., Garcia-Esteve et al., 2016; Wittkowski et al., 2010), reflecting the approach used in the original development of the PBQ (Brockington et al., 2001, 2006). However, Garcia-Esteve et al. (2016) did not find support for the original four-factor structure in either the clinical or the general sample of postpartum mothers. Also, studies differed in sample sizes, ranging from 123 (e.g., Busonera et al., 2017) to 3005 participants (Bicking Kinsey et al., 2014). Moreover, studies varied widely in the timing of PBQ administration during the first year postpartum, from as early as one month postpartum (Bicking Kinsey et al., 2014; Ohashi et al., 2016; Suetsugu et al., 2015), three months postpartum (Busonera et al., 2017; Kaneko and Honjo, 2014), to the wider range of 1 to 10 or 1 to 12 months postpartum (Anđelinović et al., 2019; Baldisserotto et al., 2023; Wittkowski et al., 2010). As some prospective studies reported better fit indices at a later PBQ administration (Göbel et al., 2024; Matsunaga et al., 2021), assessment time may be relevant and should be consistent across studies.

Studies also differed in their psychometric approach. Some conducted only exploratory factor analysis (EFA) (Bicking Kinsey et al., 2014; Busonera et al., 2017; Edhborg et al., 2011; Kaneko and Honjo, 2014; Pudasainee-Kapri et al., 2024; Suetsugu et al., 2015), while only confirmatory factor analysis (CFA) was conducted rarely (Nazaré et al., 2012). Most commonly, studies employed both EFA and CFA, typically when the original factor structure could not be adequately confirmed, and an alternative structure was subsequently explored and tested (Anđelinović et al., 2019; Baldisserotto et al., 2023; Garcia-Esteve et al., 2016; Göbel et al., 2024; Kalfon-Hakhmigari et al., 2025; Lavallée et al., 2026; Saur et al., 2023; Wittkowski et al., 2010).

Furthermore, substantial variations were observed in the final numbers of factors and retained items across countries. For example, among studies reporting a four-factor solution, the number of retained items ranged from 14 (Suetsugu et al., 2015), 24 (Göbel et al., 2024) to 25 (Edhborg et al., 2011; Garcia-Esteve et al., 2016; Palacios-Hernández et al., 2026). However, substantial variability was also observed among studies identifying a one-factor solution, with the number of retained items ranging from 10 (Bicking Kinsey et al., 2014) to 12 (Nazaré et al., 2012; Saur et al., 2023), 14 (Lavallée et al., 2026), 16 (Kaneko and Honjo, 2014), or even 22 items (Demanche et al., 2021). Furthermore, the same number of retained items was associated with different factor structures across studies. For example, 14-item versions have been reported to yield one-factor (Lavallée et al., 2026), two-factor (Kalfon-Hakhmigari et al., 2025), and four-factor (Suetsugu et al., 2015) solutions.

To summarise, given the wide variety of factor solutions and the number of items in validation studies, and previous suggestions to test the PBQ measurement invariance (Saur et al., 2023), there is a need to establish a short version of the PBQ with a consistent factor structure across countries, which would demonstrate metric invariance. To overcome some of the shortcomings of previous studies and methodological inconsistencies, we included data from multiple countries in a single study, using a common methodological approach and administering the PBQ within the same 6–12-week postpartum window across countries (Ayers, Handelzalts, & the INTERSECT Consortium, 2025). This timeframe falls within the early postpartum periods used in previous PBQ validation studies and reduces heterogeneity related to differences in assessment timing. The aim of this study was thus to establish a short version of the PBQ that retains content from all four original factors, and has adequate psychometric properties across countries, in terms of the factor structure, reliability (internal consistency and test-retest reliability), associations with depressive symptoms, and measurement invariance (configural, metric, and scalar).

Methods

Participants

The total sample consisted of 1731 postpartum women from five countries (Croatia, Iceland, Italy, Lithuania, and Switzerland) who had given birth within the past 6–12 weeks. Most participants had higher education (72.4%), were married or in a civil partnership (71.4%), and slightly more than half (53.8%) were primiparous. Detailed sociodemographic and obstetric characteristics of the study participants, both overall and by country, are presented in Table 1.

Significant cross-country differences were observed across all sociodemographic and obstetric characteristics (Table 1). Pairwise comparisons indicated that infants were slightly younger in Switzerland than in all other countries and in Lithuania than in Croatia, Iceland and Italy. Gestational age at birth was significantly lower in Lithuania than in all other countries. The proportion of women with higher education was significantly higher in Lithuania than in Italy. Iceland had a significantly higher prevalence of previous pregnancy loss than Croatia, Italy and Lithuania, and a significantly higher prevalence of current psychological or mental health problems than all other countries. All pairwise comparisons are presented in Supplementary Table S1.

Longitudinal PBQ data at 26 weeks postpartum were available from Italy and Switzerland. Accordingly, the test–retest analysis included all participants with PBQ data available at both assessment points (N = 352; Italy, N = 113; Switzerland, N = 239).

Measures

The Postpartum Bonding Questionnaire (PBQ; Brockington et al., 2001) was used to assess difficulties in mother–infant bonding. It is a self-report questionnaire designed to assess difficulties in the mother–infant relationship, consisting of 25 items rated on a 6-point Likert scale (0 - *always* to 5 - *never*), with some items reversely coded, where higher scores indicate greater bonding difficulties (See Supplementary

Table 1
Participant characteristics by country.

Sociodemographic and obstetric characteristics		Croatia (N = 380)	Iceland (N = 613)	Italy (N = 165)	Lithuania (N = 327)	Switzerland (N = 246)	p	Total (N = 1731)
Infant's age (in weeks) (Mean ± SD)		7.72 ± 1.41	7.52 ± 1.48	7.81 ± 1.79	7.06 ± 1.08	6.54 ± 1.07	<.001	7.36 ± 1.44
Gestational age at birth (Mean ± SD)		39.57 ± 1.32	39.34 ± 1.60	39.2 ± 1.66	38.6 ± 2.18	39.37 ± 1.61	<.001	39.24 ± 1.71
Maternal age (years) (%)	18–19	0.3	0.2	0.6	0.6	-	<.001	0.3
	20–24	6.1	7.2	3.6	8.0	2.4		6.1
	25–29	26.8	27.9	13.9	25.1	19.9		24.7
	30–34	37.1	36.5	40.6	36.7	53.3		39.5
	35–39	22.6	20.2	24.8	18.7	20.7		21.0
	40+	7.1	8.0	16.4	11.0	3.7		8.5
Education (%)	No formal education	-	0.5	0.6	-	-	<0.001	0.2
	Elementary/primary education	0.5	7.0	1.3	2.1	5.7		3.9
	High/secondary school education	26.5	19.9	32.5	22.6	23.2		23.5
	Higher education	73.0	72.6	65.6	75.2	71.1		72.4
Income (%)	Below average	1.6	14.6	8.1	2.4	11.4	<0.001	8.4
	Average	61.4	64.5	76.7	58.7	68.7		63.9
	Above average	37.0	20.9	15.1	38.8	19.9		27.7
Relationship status (%)	Married or in a civil partnership	75.8	77.4	41.9	78.3	59.3	<.001	71.4
	Living with a partner	23.2	16.9	38.1	20.2	37.4		23.8
	In a relationship but not living together	0.3	1.0	16.8	0.3	1.6		2.2
	Single	0.5	4.5	2.6	0.9	1.2		2.3
	Separated/divorced or other	0.3	0.2	0.6	0.3	0.4		0.3
Parity (%)	Primiparous	56.0	44.8	68.2	54.7	60.9	<.001	53.8
	Multiparous	44.0	55.2	31.8	45.3	39.1		46.2
Birth method (%)	Vaginal birth	77.9	78.4	76.1	70.0	65.4	<0.001	74.6
	Assisted vaginal birth, e.g. forceps, ventouse	0.3	8.6	7.4	1.2	12.2		5.7
	Emergency caesarean	13.7	8.1	9.8	17.7	14.2		12.2
	Elective caesarean	8.2	4.9	6.7	11.0	8.1		7.4
Previous pregnancy loss (%)	Yes	21.6	32.9	19.6	26.6	-	<0.001	72.8
	No	78.4	67.1	80.4	73.4	-		27.2
Current psychological or mental health problem (%)	Yes	2.1	21.9	3.9	0.6	3.7	<0.001	9.2
	No	91.8	67.2	90.2	88.4	94.3		82.7
	I don't know	6.1	10.9	5.9	11.0	2.0		8.1

material Table S2). The original scale comprises four subscales: impaired bonding, which captures general difficulties in the mother–infant relationship, including emotional distance and reduced affection or enjoyment; rejection and pathological anger, which reflects rejection, resentment, and anger toward the infant; infant-focused anxiety, which assesses anxiety and concerns centred on the infant and their care; and risk of abuse, which captures thoughts or behaviours indicating a risk of harming the infant. Validated versions of the PBQ were used in Croatia (Anđelinović et al., 2019), Italy (Busonera et al., 2017), and Switzerland (French; Demanche et al., 2021), whereas in Lithuania and Iceland, the instrument was translated following a forward–backward translation procedure.

Depressive symptoms were assessed using the Edinburgh Postnatal Depression Scale (EPDS; Cox et al., 1987), a 10-item self-report measure of postpartum depressive symptoms. Each item is rated on a 4-point scale (0 to 3), with higher scores indicating greater depressive symptoms. Validated versions of the EPDS were used in all participating countries.

The full sociodemographic and obstetric questionnaire used in the study can be found on the INTERSECT study website https://www.intersectstudy.org/

Procedure

This study used data collected as part of the INTERSECT study from countries that included the full version of the PBQ in the INTERSECT survey. All participants who completed at least 80% of the PBQ items were included in the analysis. For a full description of the INTERSECT study, see Handelzalts et al. (2025).

The 6–12-week postpartum assessment window was predefined as part of the broader INTERSECT study, which focused on maternal mental health during the early postpartum period, and was not selected specifically for the present PBQ analysis. Nevertheless, applying the same relatively narrow assessment window across all participating countries enabled the PBQ factor structure to be examined while minimizing variation associated with differences in the timing of questionnaire administration.

Ethical approval was obtained separately in all participating countries (Ethics Committee of the Catholic University of Croatia: Class: 641-03/21-03/21; No: 498-16/2-22-04; The National Bioethics Committee in Iceland (VSN-21-237); Research Ethics Committees of the Catholic University of Milan [Reference: 05-22] and S. Giuseppe Hospital in Milan [Reference: 550/2022]; Kaunas Regional Biomedical Research Ethics Committee in Lithuania [Reference: BE-2-117; 2021-10-14]; Commission Cantonale d'Éthique de la Recherche sur l'Être Humain in Switzerland [Reference: 2022-00284]).

The study followed the same inclusion criteria, recruitment procedures, and survey measures across all countries as described elsewhere (Ayers et al., 2021). Participants were recruited during pregnancy or postpartum through routine maternity care services, including hospitals, clinics, and birth centers, and not through social media. Women recruited during pregnancy who agreed to participate provided their contact details and were subsequently invited to complete the online survey during the predefined 6–12-week postpartum period. Women recruited postpartum were invited to complete the same survey within this period. Access to the online survey varied slightly by country. In Lithuania, the survey link was sent by email or SMS, according to participants' contact preferences, and in Croatia by email or WhatsApp, using contact details provided during the initial hospital-based assessment before birth. In Switzerland, participants received an email invitation containing a personal link to the questionnaire. In Italy, participants initially received study information and a QR-code flyer during antenatal classes or while attending scheduled medical appointments; those who consented to follow-up subsequently received the second-phase questionnaire, which included the PBQ, via a direct email link. In Iceland, email invitations were sent using addresses obtained from the National Birth Registry.

Informed consent was obtained from all participants prior to participation. Data were collected using online questionnaires between October 2021 and December 2023. The data were processed and coded using a common coding framework to ensure consistency across countries. Anonymised datasets were subsequently transferred to the data server at City St. George's, University of London. A data protection impact assessment was completed and approved, and appropriate research governance and data-sharing agreements were put in place between the coordinating and collaborating institutions.

Data analysis

Data were analysed using IBM SPSS Statistics version 27 and Mplus version 8.6. Descriptive statistics were calculated for sociodemographic and obstetric characteristics and are presented as means and standard deviations for continuous variables and percentages for categorical variables. Cross-country differences were examined using one-way analysis of variance (ANOVA) for normally distributed continuous variables, the Kruskal–Wallis test for non-normally distributed continuous variables, and Pearson's chi-square or Fisher's exact tests, as appropriate, for categorical variables. Missing data on PBQ items were minimal, ranging from 0% to 0.69% across items, with an overall item-level missingness of 0.11%. Missing responses were handled using the default WLSMV procedure in Mplus for categorical indicators. Items were recoded as necessary prior to the psychometric analyses so that higher scores consistently reflected greater bonding difficulties.

Prior to EFA, an initial item-screening procedure was conducted to identify items showing consistently poor psychometric functioning across countries. Response distributions (skewness and kurtosis), standard deviations, inter-item correlations, and corrected item–total correlations (CITCs) were examined in the full sample and separately within each country. This preliminary screening was intended to identify clearly poorly performing items for potential exclusion before examination of the latent structure, in line with the aim of developing a concise version of the PBQ. Item exclusion at this stage was based on the overall pattern of evidence across multiple indicators and countries rather than on a single psychometric criterion.

Given the ordinal nature of the items and the use of a non-clinical sample, skewness, kurtosis, and standard deviations were interpreted descriptively rather than used as strict exclusion criteria. As the PBQ is designed to assess impairments in the mother–infant relationship, low frequencies of endorsement of higher response categories and restricted variability were expected. Applying strict distributional criteria would therefore have risked excluding items reflecting less frequently endorsed but potentially clinically relevant bonding difficulties.

Greater emphasis was therefore placed on item relationships, specifically inter-item correlations and CITCs. CITCs below .20 and mean inter-item correlations below .10 were considered indicators of weak associations with the overall construct rather than strict stand-alone exclusion thresholds. To distinguish consistently weak item functioning from country-specific variation, poor performance was required to occur in at least three of the five countries, representing the majority of country samples. This study-specific majority-of-countries criterion was intended to reduce the likelihood of excluding items because of weak performance in only one or two countries. The use of inter-item and corrected item–total correlations for preliminary item evaluation was informed by established recommendations for scale development (e.g., Boateng et al., 2018; Streiner et al., 2015; Zijlmans et al., 2018), while the thresholds were selected to identify clearly weak item relationships at this preliminary screening stage. Item exclusion was based on the overall pattern of evidence across psychometric indicators and countries. A stratified random split by country was performed, with participants within each country randomly allocated with roughly equal probability to either the exploratory factor analysis or confirmatory factor analysis subsample (development sample, N = 857; validation

sample, N = 874). This approach ensured that both datasets had similar cross-country distributions (see Supplementary Material Table S3).

Exploratory factor analysis (EFA) with oblique Geomin rotation, specifying one- to five-factor solutions, was conducted on the development sample to examine the factor structure of the remaining item pool. This range was selected to reflect the original multidimensional structure of the PBQ (four factors) and prior findings indicating up to five factors, while also allowing for a more parsimonious structure suitable for a short version. Item retention decisions were based on factor loadings, cross-loadings, item–total correlations, and conceptual interpretability. Items were retained if they demonstrated a primary loading ≥ .40, a difference of at least .20 between primary and secondary loadings, and clear conceptual relevance to the construct (Tabachnick and Fidell, 2019; Costello and Osborne, 2005). Competing factor solutions were evaluated based on statistical criteria and conceptual interpretability. Once the most interpretable solution was identified, poorly performing items were removed iteratively, with the factor structure re-evaluated following item removal. The number of factors was not fixed during this process but was re-evaluated as the item pool was reduced. Factors were required to contain at least three items to ensure adequate stability and interpretability of the factor structure, as factors with fewer than three items are generally considered weak and unstable (Costello and Osborne, 2005).

Confirmatory factor analysis (CFA) using WLSMV estimation was then performed on the validation sample to evaluate this structure. Model fit was assessed using chi-square statistic (χ^2, $p \geq .05$), comparative fit index (CFI ≥ .90), Tucker–Lewis index (TLI ≥ .90), root mean square error of approximation (RMSEA ≤ .08), standardized root mean square residual (SRMR ≤ .08), relative chi-square ($\chi^2/df \leq 3$), and item loadings (≥ .40), following recommended criteria (Tabachnick and Fidell, 2019).

Measurement invariance across countries was examined using multigroup CFA with categorical indicators. Following established recommendations (Van de Schoot et al., 2012), measurement invariance was evaluated sequentially at the configural, metric, and scalar levels. Configural invariance assessed whether the same factor structure was present across countries. Metric invariance evaluated the equivalence of factor loadings across groups, whereas scalar invariance examined the equivalence of item thresholds across countries, allowing assessment of whether observed scores could be interpreted similarly across groups. Prior to invariance testing, examination of item response distributions revealed sparse and empty response categories for several items across countries. Response categories were therefore collapsed to facilitate threshold estimation in multi-group CFA with categorical indicators while preserving the ordinal nature of the data. Robust weighted least squares estimation (WLSMV) was used throughout the analyses.

Model fit indices, internal consistency estimates, and correlations between the long and short versions were also examined. Hypothesis testing for construct validity was based on the association with depressive symptoms. As postpartum depression and bonding difficulties are conceptually and empirically related but distinct constructs, a positive but not strong correlation was expected (Edhborg et al., 2011). Associations between the long and short versions, as well as with depressive symptoms, were examined using Spearman's rank correlations due to non-normal and skewed PBQ score distributions. Effect size interpretation followed Cohen's conventions, where r =.10, .30, and .50 represent small, medium, and large effects, respectively (Cohen, 1988).

Test–retest reliability was assessed using the intraclass correlation coefficient (ICC) in a subsample from Italy and Switzerland, as well as within individual countries. Internal consistency was evaluated using Cronbach's α and McDonald's ω. Values of .70 and above were considered acceptable, .80 and above good, and .90 and above excellent, based on commonly recommended guidelines (DeVellis, 2017).

Results

Exploratory factor analysis

Initial item screening identified Items 9, 16, and 18 as showing consistently poor psychometric performance across multiple indicators and country samples. Specifically, all three items showed weak inter-item relationships in at least three countries, while Items 16 and 18 also showed low corrected item–total correlations; their response distributions and limited variability provided additional evidence considered during the screening process. Based on this converging evidence, these three items were removed prior to EFA, which was conducted on the remaining 22 items. Model fit indices and factor loadings for the initial one- to five-factor solutions are presented in Supplementary Tables S4 and S5. The three-, four-, and five-factor solutions were not retained due to poorly defined factors, including substantial cross-loadings and factors represented by fewer than three items. Across the two- to five-factor solutions, Items 12 ("My baby cries too much") and 25 ("My baby is easily comforted") consistently loaded on a separate factor. Although this factor was stable across solutions, it reflected infant behavioural characteristics rather than maternal bonding; therefore, these items were considered conceptually distinct from the construct of interest and were removed.

The two-factor solution was selected as the most interpretable starting point for item reduction. Following iterative removal of cross-loading items (Items 7, 8, 13, 14, 15, 19, 21, and 22), with the factor structure re-evaluated after each item removal, the remaining items ultimately supported a stable and interpretable unidimensional solution. Thus, the one-factor structure was not specified a priori but emerged during the iterative item-reduction process. The final one-factor model comprised 12 items and demonstrated good model fit ($\chi^2(54) = 228.03$, $p < 0.001$; RMSEA = .061, 90% CI [.053–.070], CFI = .965, TLI = .957, SRMR = .073). Factor loadings for the final model are presented in Table 2.

Confirmatory factor analysis

The one-factor structure identified in the EFA was subsequently evaluated using CFA in the validation subsample. The 12-item model demonstrated good model fit ($\chi^2(54) = 277.35$, $p < 0.001$; RMSEA = .069 [90% CI .061–.077], CFI = .965, TLI = .958, SRMR = .065). All items loaded significantly on the latent factor, with standardized factor loadings ranging from .43 to .85. Factor loadings and model estimates are presented in Table 2.

Table 2
Factor loadings for the one-factor PBQ 12-item short version derived from exploratory factor analysis (EFA) and tested with confirmatory factor analysis (CFA).

Item	Content	EFA Loading	CFA Loading
1.	I feel close to my baby.	.822	.848
2.	I wish the old days when I had no baby would come back.	.700	.686
3.	I feel distant from my baby.	.791	.800
4.	I love to cuddle my baby.	.812	.791
5.	I regret having this baby.	.745	.756
6.	The baby does not seem to be mine.	.674	.645
10.	I love my baby to bits.	.726	.789
11.	I enjoy playing with my baby.	.773	.767
17.	I wish my baby would somehow go away.	.782	.735
20.	I am afraid of my baby.	.694	.721
23.	I feel the only solution is for someone else to look after my baby.	.681	.706
24.	I feel like hurting my baby.	.469	.432

Measurement invariance

Due to limited sample sizes in some country subsamples within the validation split, measurement invariance analyses were conducted using the full sample. Examination of item response distributions revealed sparse or absent response categories in several countries (Supplementary Table S6). To facilitate estimation in multi-group CFA with categorical indicators, response categories were collapsed while preserving the ordinal nature of the data. Despite this adjustment, Items 6 and 17 exhibited extremely low variability in Italy, with responses concentrated almost entirely within a single category. Because these items did not provide sufficient information for stable parameter estimation across countries, they were excluded from further cross-country evaluation, resulting in a 10-item version used for measurement invariance testing.

Measurement invariance across countries was examined using multi-group CFA with categorical indicators. The configural model demonstrated acceptable fit, supporting the presence of the same one-factor structure across countries, $\chi^2(175) = 446.70$, $p < 0.001$; RMSEA = .067, 90% CI [.059, .075], CFI = .961, TLI = .950, SRMR = .080. Metric invariance was subsequently tested by constraining factor loadings to equality across groups. The metric model also demonstrated acceptable fit, $\chi^2(211) = 555.45$, $p < 0.001$; RMSEA = .069, 90% CI [.062, .076], CFI = .951, TLI = .948, SRMR = .088. Although the chi-square difference test between the configural and metric models was statistically significant, $\Delta\chi^2(36) = 128.04$, $p < .001$, changes in approximate fit indices were minimal (ΔCFI = .010; ΔRMSEA = .002), supporting metric invariance.

Scalar invariance was then examined by additionally constraining item thresholds to equality across countries. The scalar model demonstrated acceptable fit, $\chi^2(247) = 556.96$, $p < 0.001$; RMSEA = .060, 90% CI [.054, .067], CFI = .956, TLI = .960, SRMR = .089. Comparison of the scalar and metric models indicated no significant deterioration in model fit, $\Delta\chi^2(36) = 37.76$, $p = .389$. Together, these findings support configural, metric, and scalar invariance of the 10-item PBQ short version across the participating countries.

Following the removal of Items 6 and 17 during measurement invariance testing, the factor structure of the resulting 10-item version was re-evaluated using CFA in the validation sample. The unidimensional model showed good fit (CFI = .973, TLI = .965, SRMR = .052, RMSEA = .072, 90% CI [.063, .082]), with factor loadings presented in Table 3.

Other psychometric parameters

Internal consistency of the scale. Internal consistency of the final 10-item version was good, with Cronbach's α = .81 and McDonald's ω = .82. The 25-item full version demonstrated slightly higher reliability (α = .87, ω = .89), reflecting a modest reduction in internal consistency in the shortened version.

Test-retest reliability. The 10-item short version demonstrated an intraclass correlation coefficient (ICC; two-way mixed, absolute agreement) of .73 (95% CI [.64, .79], $p < .001$) in the test–retest analysis.

Country-specific analyses revealed some variability in temporal stability. In Italy, test–retest reliability was moderate (ICC = .56, 95% CI [.42, .68], $p < .001$; N = 113), whereas in Switzerland it was good (ICC = .78, 95% CI [.69, .84], $p < .001$; N = 239).

Long and short version comparison and correlation. In full sample, The 10-item version yielded CFI = .955, TLI = .942, SRMR = .047, and RMSEA = .076 (90% CI [.069, .083]). The corresponding values for the 25-item version were CFI = .899, TLI = .890, SRMR = .068, and RMSEA = .073 (90% CI [.070, .075]). The two versions were strongly correlated ($\rho = .80$, $p < .001$).

Association with depressive symptoms. The 10-item PBQ short version showed a moderate positive association with depressive symptoms measured by the EPDS ($\rho = .37$, $p < .001$), as hypothesized.

Table 3
Factor loadings for the final one-factor PBQ 10-item short version.

Item	Content	Factor loading
1.	I feel close to my baby.	.854
2.	I wish the old days when I had no baby would come back.	.684
3.	I feel distant from my baby.	.803
4.	I love to cuddle my baby.	.803
5.	I regret having this baby.	.750
10.	I love my baby to bits.	.794
11.	I enjoy playing with my baby.	.782
20.	I am afraid of my baby.	.680
23.	I feel the only solution is for someone else to look after my baby.	.665
24.	I feel like hurting my baby.	.404

Discussion

The present study developed a short version of the PBQ using a large multi-country sample and a common measurement framework. Exploratory and confirmatory analyses initially supported a 12-item unidimensional structure with good psychometric properties across countries. Two items were subsequently removed during measurement invariance testing due to low response variability in one country, resulting in a 10-item unidimensional version without affecting the structure or interpretation of the measure. Importantly, the retained items preserved content coverage across all four conceptual domains of the original PBQ bonding, rejection, anxiety, and thoughts of harming the infant, with at least one item retained from each original factor. At the same time, the unidimensional structure supports the conceptualization of maternal–infant bonding difficulties as a general rather than multidimensional construct. The 10-item version showed strong correspondence with the full scale, improved model fit, good internal consistency, temporal stability, and the expected association with depressive symptoms. Furthermore, evidence of configural, metric, and scalar invariance across countries supports its use for cross-country comparisons.

The present findings align with growing evidence supporting unidimensional PBQ solutions across diverse contexts (e.g., Bicking-Kinsey et al., 2014; Bleijenbergh et al., 2022; Kaneko and Honjo, 2014; Nazaré et al., 2012), with recent studies increasingly converging on reduced unidimensional versions (Lavallée et al., 2026; Saur et al., 2023). Unlike previous studies, which generally examined the PBQ within individual countries and produced varying factor structures and item compositions, the present study identified a common 10-item unidimensional structure across multiple countries and languages. Moreover, the demonstration of configural, metric, and scalar invariance indicates that the short version measures bonding difficulties comparably across these contexts, supporting cross-country comparisons that could not be established from previous country-specific findings.

This may reflect the fact that bonding-related difficulties are expressed through heterogeneous indicators that capture different manifestations of a shared underlying dimension (Reise, 2012; Rodriguez et al., 2016). When items retained across validation studies are considered in relation to the original PBQ domains, a nuanced pattern emerges: items reflecting impaired bonding and rejection are more consistently retained, whereas anxiety-related items appear less stable across factor structures. In addition, the abuse factor, typically represented by only a small number of items, has shown less consistent replication across studies (e.g., Garcia-Esteve et al., 2016; Ohashi et al., 2016; Palacios-Hernández et al., 2026). Although these patterns are not uniform, they suggest that PBQ domains vary in their stability across contexts. Taken together, the findings suggest that variability across studies may reflect differences in the expression and measurement of a shared underlying construct rather than the absence of a coherent

structure. In this context, the convergence toward a unidimensional solution across countries under a common measurement framework supports interpreting the PBQ as capturing a broad dimension of bonding difficulties.

These findings have important implications for conceptualizing maternal–infant bonding difficulties. Although the original PBQ was structured around multiple domains, the convergence toward a unidimensional solution suggests that these domains may not function as clearly separable components. Instead, PBQ scores appear to reflect a broad dimension of bonding difficulties encompassing a range of affective responses toward the infant (Nakić Radoš et al., 2024), including emotional distance, irritation, and rejection. These responses may co-occur and fluctuate within individuals rather than represent distinct dimensions. This interpretation is consistent with theoretical accounts of bonding as a multifaceted but integrated process (Brockington, 2004; Nakić Radoš et al., 2024) and with broader dimensional approaches to psychopathology and relational functioning (Kotov et al., 2017). Such a perspective may help explain the inconsistent multidimensional structures reported across studies.

Previous attempts to test PBQ invariance showed configural, metric, and scalar invariance across different time assessments (5 days vs. 1 month postpartum) in Japan (Matsunaga et al., 2021) and scalar invariance in samples of mothers with and without depressive symptoms in Brazil (Saur et al., 2023). However, to the authors' best knowledge, the only attempt to test invariance across different languages, but still within the same country, was a study in the USA, where weak metric invariance was observed for English and Spanish administrations of the PBQ (Lavallée et al., 2026). Therefore, by demonstrating configural, metric, and scalar invariance across countries and languages, the present study further supports the use of the 10-item PBQ short form.

Item retention in the present study reflected a balance between statistical parsimony and conceptual coverage. Although some items showed lower loadings or less consistent retention across prior validation studies, they were retained due to their theoretical and clinical relevance. For example, harm-related items, often excluded due to low endorsement or statistical challenges, may capture clinically meaningful experiences not represented by more frequently endorsed indicators (Brockington, 2004). Retaining items from multiple content domains allowed the scale to capture a broader range of bonding-related experiences, consistent with recommendations emphasizing construct representation alongside internal consistency (Clark and Watson, 1995; Haynes et al., 1995). This may be particularly important for multifaceted, context-dependent constructs, where reliance on statistical criteria alone could exclude clinically meaningful content.

Beyond its structural validity and measurement invariance, the 10-item PBQ also demonstrated satisfactory psychometric properties. Internal consistency was good, and test–retest reliability indicated acceptable temporal stability. Furthermore, the strong correlation with the original 25-item version suggests that the short form retains a substantial proportion of the information captured by the full scale while reducing respondent burden. The moderate association with postpartum depressive symptoms was consistent with the hypothesized relationship between postpartum depression and bonding difficulties, providing additional support for construct validity.

Strengths and limitations

The present study used a common measurement framework across multiple countries, and the standardized assessment window minimized variability associated with changes in maternal experiences across the postpartum period. In addition, the use of exploratory and confirmatory analyses provided a rigorous evaluation of the measure's structure (Brown, 2015; Wittkowski et al., 2020). A further strength of the study was the demonstration of configural, metric, and scalar invariance across countries, supporting the use of the scale for cross-country comparisons. Finally, the large and culturally diverse sample strengthened the robustness of the findings and supported the convergence toward a unidimensional structure across cultural contexts.

Several limitations should be considered. First, the study relied on community samples in which most participants were women of relatively high socioeconomic status and in stable relationships, potentially limiting generalizability to clinical populations, where bonding difficulties may present differently (Brockington et al., 2001). Second, national samples differed significantly in their sociodemographic and obstetric characteristics, which may limit the generalizability of the findings despite evidence of measurement invariance across countries. Third, hypothesis testing for construct validity was limited to the association with depressive symptoms. Future studies should include additional measures of bonding and related constructs, including anxiety, parenting stress, maternal attachment, and observational measures of parent–infant interaction, to provide a more comprehensive evaluation of convergent and discriminant validity. Moreover, country-specific sample sizes were unequal, with a comparatively smaller sample from Italy, which may have affected the precision of parameter estimates in the multi-group analyses. Although measurement invariance was supported, these findings should therefore be interpreted with some caution and replicated in studies with larger and more balanced country samples. Finally, assessment was limited to the 6–12 week postpartum period, although bonding-related experiences may change over time and influence the structure of the measure (Wittkowski et al., 2020).

Implications for practice and research

Clinically, the findings support the use of the PBQ as a screening tool for identifying general bonding difficulties. The 10-item version may be particularly useful in routine postpartum care, large survey batteries, repeated assessments, and cross-national research, where reducing assessment burden is important. Its items represent all four content domains of the original PBQ, while its unidimensional structure and demonstrated measurement invariance support the assessment and comparison of general bonding difficulties across countries.

In addition, the convergence toward a unidimensional structure suggests that overall levels of bonding difficulties may be more informative than specific symptom profiles, particularly in early identification contexts. However, as with other self-report screening instruments, elevated PBQ scores should not be considered sufficient for establishing clinically significant bonding disorders. When severe bonding difficulties are suspected, follow-up assessment using a structured clinical interview, such as the Stafford Interview (Brockington et al., 2017), remains essential and represents the current gold standard for clinical evaluation.

From a research perspective, these findings highlight the need for caution in interpreting the PBQ as a multidimensional measure. The variability in factor structures and item composition across studies suggests that subscale scores may not be stable or comparable across contexts. Accordingly, future research may benefit from focusing on the PBQ as a dimensional measure of general bonding impairment, while further examining how specific bonding-related experiences are expressed across populations, cultures, and stages of the postpartum period.

Conclusions

In summary, the present study demonstrates that the PBQ converges toward a unidimensional structure when examined within a large, multicountry sample while using similar inclusion criteria and assessment time point. The resulting short 10-item version provides a parsimonious and psychometrically robust assessment while retaining broad conceptual coverage. These findings support the interpretation of the PBQ as a measure of general bonding difficulties and contribute to clarifying its structure in the context of existing inconsistent findings across the literature.

Ethics declaration

Written informed consent to take part in the study and to publish the article has been obtained from all participants or their legal representatives. The privacy rights of participants have been observed.

This study was performed in compliance with relevant laws, regulatory frameworks and guidelines where the research took place. This study was approved by the Kaunas Regional Biomedical Research Ethics Committee in Lithuania. (Approval No. BE-2-117; 2021-10-14)

CRediT authorship contribution statement

Gabija Jarašiūnaitė-Fedosejeva: Writing – review & editing, Writing – original draft, Supervision, Methodology, Investigation, Formal analysis, Conceptualization. **Sandra Nakić Radoš:** Writing – review & editing, Writing – original draft, Supervision, Methodology, Investigation, Conceptualization. **Susan Ayers:** Writing – review & editing, Visualization, Resources, Conceptualization. **Maja Brekalo:** Writing – review & editing, Investigation. **Raquel Costa:** Writing – review & editing, Resources, Investigation. **Antje Horsch:** Writing – review & editing, Resources, Investigation. **Chiara Ionio:** Writing – review & editing, Resources, Investigation. **Maor Kalfon-Hakhmigari:** Writing – review & editing, Investigation. **Tiago Miguel Pinto:** Writing – review & editing, Resources, Investigation. **Valentine Rattaz:** Writing – review & editing, Investigation. **Olga Riklikienė:** Writing – review & editing, Resources, Investigation. **Valgerður Lísa Sigurðardóttir:** Writing – review & editing, Resources, Investigation. **Emma Marie Swift:** Writing – review & editing, Investigation. **Rebecca Webb:** Writing – review & editing, Investigation. **Jonathan Eliahu Handelzalts:** Writing – review & editing, Writing – original draft, Visualization, Supervision, Resources, Investigation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

None.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.midw.2026.105020.

References

Anđelinović, M., Nakić Radoš, S., Matijaš, M., 2019. Validation of Postpartum Bonding Questionnaire in a sample of Croatian mothers and fathers. In: Pačić-Turk, L., Knežević, M. (Eds.), Proceedings of the 2nd International Scientific Conference Brain Mind: Promoting Individual Community Well-Being. Catholic University of Croatia, Zagreb. Book of Abstract (p. 153)Available online at https://www.bib.irb.hr/1039400.

Ayers, S., Handelzalts, J., Webb, R., INTERSECT Consortium, 2021. International survey of childbirth-related trauma (INTERSECT) (research registry no. 6439). Research Registry. Retrieved April 3, 2026, from. https://www.researchregistry.com/browse-the-registry/#home/registrationdetails/5ffc7453702012001b80a58c/.

Ayers, S., Handelzalts, J., The INTERSECT Consortium. (2025). INTERSECT: international survey of childbirth-related trauma, 2021–2024. [data collection]. 2nd Edition. UK Data Service. SN: 9295, 10.5255/UKDA-SN-9295-2.

Baldisserotto, M.L., Griep, R.H., Oates, J., Theme-Filha, M.M., 2023. Cross-cultural adaptation of the Postpartum Bonding Questionnaire (PBQ) to Brazil: assessment of the measurement properties of construct validity and reliability. Curr. Psychol. 42 (30), 26275–26285. https://doi.org/10.1007/s12144-022-03742-3.

Bicking Kinsey, C., Baptiste-Roberts, K., Zhu, J., Kjerulff, K.H., 2014. Birth-related, psychosocial, and emotional correlates of positive maternal-infant bonding in a cohort of first-time mothers. Midwifery 30 (5), e188–e194. https://doi.org/10.1016/j.midw.2014.02.006.

Bicking Kinsey, C., Hupcey, J.E, 2013. State of the science of maternal-infant bonding: a principle-based concept analysis. Midwifery 29 (12), 1314–1320. https://doi.org/10.1016/j.midw.2012.12.019.

Bleijenbergh, R., Van den Branden, L., van Gils, Y., Van de Craen, N., Van Leugenhaege, L., Kuipers, Y.J., 2022. Validation of the Postpartum Bonding Questionnaire: a cross-sectional study among Flemish mothers. Midwifery 107, 103280. https://doi.org/10.1016/j.midw.2022.103280.

Boateng, G.O., Neilands, T.B., Frongillo, E.A., Melgar-Quiñonez, H.R., Young, S.L., 2018. Best practices for developing and validating scales for health, social, and behavioral research: a primer. Front. Public Health 6, 149. https://doi.org/10.3389/fpubh.2018.00149. Article.

Brockington, I., Chandra, P., Bramante, A., Dubow, H., Fakher, W., Garcia-Esteve, L., Shieh, P.L., 2017. The Stafford Interview: a comprehensive interview for mother-infant psychiatry. Arch. Womens Ment. Health 20 (1), 107–112. https://doi.org/10.1007/s00737-016-0683-8.

Brockington, I.F., 2004. Postpartum psychiatric disorders. Lancet 363 (9405), 303–310. https://doi.org/10.1016/S0140-6736(03)15390-1.

Brockington, I.F., Fraser, C., Wilson, D., 2006. The Postpartum Bonding Questionnaire: a validation. Arch. Womens Ment. Health 9 (5), 233–242. https://doi.org/10.1007/s00737-006-0132-1.

Brockington, I.F., Oates, J., George, S., Turner, D., Vostanis, P., Sullivan, M., Loh, C., Murdoch, C., 2001. A screening questionnaire for mother-infant bonding disorders. Arch. Womens Ment. Health 3, 133–140. https://doi.org/10.1007/s007370170010.

Brown, T.A., 2015. Confirmatory Factor Analysis for Applied Research, 2nd ed. Guilford Press.

Busonera, A., Cataudella, S., Lampis, J., Tommasi, M., Zavattini, G.C., 2017. Psychometric properties of the Postpartum Bonding Questionnaire and correlates of mother-infant bonding impairment in Italian new mothers. Midwifery 55, 15–22. https://doi.org/10.1016/j.midw.2017.08.011.

Clark, L.A., Watson, D., 1995. Constructing validity: basic issues in objective scale development. Psychol. Assess. 7 (3), 309–319. https://doi.org/10.1037/1040-3590.7.3.309.

Cohen, J., 1988. Statistical Power Analysis For the Behavioral Sciences, 2nd ed. Lawrence Erlbaum Associates.

Costello, A.B., Osborne, J.W., 2005. Best practices in exploratory factor analysis: four recommendations for getting the most from your analysis. Pract. Assess. Res. Eval. 10 (7), 1–9. https://doi.org/10.7275/jyj1-4868.

Cox, J.L., Holden, J.M., Sagovsky, R., 1987. Detection of postnatal depression: development of the 10-item Edinburgh postnatal depression scale. Br. J. Psychiatry 150, 782–786. https://doi.org/10.1192/bjp.150.6.782.

Deans, C.L., 2020. Maternal sensitivity, its relationship with child outcomes, and interventions that address it: a systematic literature review. Early Child Dev. Care 190 (2), 252–275. https://doi.org/10.1080/03004430.2018.1465415.

Demanche, A., Michel, L., Chabbert, M., Wendland, J., 2022. Version française du Postpartum Bonding Questionnaire. Devenir 33 (4), 287–310. https://doi.org/10.3917/dev.214.0287.

DeVellis, R., 2017. Scale Development: Theory and Applications, 4 ed. Sage Publications.

Dişsiz, M., Bingöl, F.B., Bal, M.D., Yılmaz, Z.D.K., Karakoç, A., Bilgin, Z., 2024. The Turkish version of the Postpartum Bonding Questionnaire (PBQ): examination of the validity and reliability and scale structure. J. Pediatr. Nurs. 77, 131–139. https://doi.org/10.1016/j.pedn.2024.03.012.

Edhborg, M., Nasreen, H.E., Kabir, Z.N., 2011. Impact of postpartum depressive and anxiety symptoms on mothers' emotional tie to their infants 2–3 months postpartum: a population-based study from rural Bangladesh. Arch. Womens Ment. Health 14 (4), 307–316. https://doi.org/10.1007/s00737-011-0221-7.

Garcia-Esteve, L., Torres, A., Lasheras, G., Palacios-Hernández, B., Farré-Sender, B., Subirà, S., Brockington, I.F., 2016. Assessment of psychometric properties of the Postpartum Bonding Questionnaire (PBQ) in Spanish mothers. Arch. Womens Ment. Health 19 (2), 385–394. https://doi.org/10.1007/s00737-015-0589-x.

Göbel, A., Lüersen, L., Asselmann, E., Arck, P., Diemert, A., Garthus-Niegel, S., Martini, J., 2024. Psychometric properties of the maternal postnatal attachment scale and the Postpartum Bonding Questionnaire in three German samples. BMC Pregnancy ChildBirth 24 (1), 789. https://doi.org/10.1186/s12884-024-06964-4.

Handelzalts, J.E., Ayers, S., Webb, R., Constantinou, G., Lucas, G., Grollman, C., Žutić, M., 2025. Cross-national risk factors for childbirth-related PTSD: findings from the INTERSECT study. Psychol. Med. 55, e349. https://doi.org/10.1017/S0033291725102298.

Haynes, S.N., Richard, D.C.S., Kubany, E.S., 1995. Content validity in psychological assessment: a functional approach to concepts and methods. Psychol. Assess. 7 (3), 238–247. https://doi.org/10.1037/1040-3590.7.3.238.

Johnson, K., 2013. Maternal-infant bonding: a review of literature. Int. J. Childbirth Educ. 28 (3), 17–22.

Kalfon-Hakhmigari, M., Segal, H., Peled, Y., Handelzalts, J.E., 2025. Hebrew validation of the Postpartum Bonding Questionnaire: a study of mothers and fathers. J. Reprod. Infant. Psychol. 43 (2), 459–471. https://doi.org/10.1080/02646838.2023.2247014.

Kaneko, H., Honjo, S., 2014. The psychometric properties and factor structure of the Postpartum Bonding Questionnaire in Japanese mothers. Psychology 05 (9), 1135–1142. https://doi.org/10.4236/psych.2014.59126.

Kotov, R., Krueger, R.F., Watson, D., Achenbach, T.M., Althoff, R.R., Bagby, R.M., Brown, T.A., Carpenter, W.T., Caspi, A., Clark, L.A., Eaton, N.R., Forbes, M.K., Forbush, K.T., Goldberg, D., Hasin, D., Hyman, S.E., Ivanova, M.Y., Lynam, D.R., Markon, K., Zimmerman, M., 2017. The hierarchical taxonomy of psychopathology (HiTOP): a dimensional alternative to traditional nosologies. J. Abnorm. Psychol. 126 (4), 454–477. https://doi.org/10.1037/abn0000258.

Lavallée, A., Warmingham, J.M., Reimers, M.A., Curtin, P., Kyle, M.H., Austin, J., Dumitriu, D., 2026. Development of a revised and abbreviated version of the Postpartum Bonding Questionnaire (PBQ-R): first US validation and association to child outcomes. Infant. Ment. Health J. 47 (1), e70052. https://doi.org/10.1002/imhj.70052.

Le Bas, G., Youssef, G., Macdonald, J.A., Teague, S., Mattick, R., Honan, I., Hutchinson, D., 2022. The role of antenatal and postnatal maternal bonding in infant development. J. Am. Acad. Child Adolesc. Psychiatry 61 (6), 820–829.e1. https://doi.org/10.1016/j.jaac.2021.08.024.

Matsunaga, A., Ohashi, Y., Sakanashi, K., Kitamura, T., 2021. Factor structure of the Postpartum Bonding Questionnaire: configural invariance and measurement invariance across postpartum time periods. J. Psychiatr. Res. 135, 1–7. https://doi.org/10.1016/j.jpsychires.2020.11.017.

Montealegre-Ramón, M.D.P., Martínez-Fuentes, M.T., Pérez-López, J., Sierra-García, P., 2024. Maternal sensitivity, early interactions and infant attachment: a systematic review. Early Child Dev. Care 194 (11–12), 1093–1108. https://doi.org/10.1080/03004430.2024.2386324.

Nakić Radoš, S., Hairston, I., Handelzalts, J., 2024. The concept analysis of parent-infant bonding during pregnancy and infancy: a systematic review and meta-synthesis. J. Reprod. Infant. Psychol. 42 (2), 142–165. https://doi.org/10.1080/02646838.2022.2162487.

Nazaré, B., Fonseca, A., Canavarro, M.C., 2012. Avaliação da ligação parental ao bebé após o nascimento: análise factorial confirmatória da versão portuguesa do Postpartum Bonding Questionnaire (PBQ). Anál. Psicol. 10 (1), 47–61. http://hdl.handle.net/10400.12/3442.

Ohashi, Y., Kitamura, T., Sakanashi, K., Tanaka, T., 2016. Postpartum bonding disorder: factor structure, validity, reliability and a model comparison of the Postnatal Bonding Questionnaire in Japanese mothers of infants. Healthcare 4 (3), 50. https://doi.org/10.3390/healthcare4030050.

Palacios-Hernández, B., Cruz-Alaniz, Y., Austria-Corrales, F., Sánchez-Aviléz, C., Martínez-Rodríguez, M.S., Téllez-Alanís, B., Harmony, T., 2026. Psychometric properties and validation of the Postpartum Bonding Questionnaire (PBQ) in Mexican mothers. Arch. Womens Ment. Health 29 (3), 67. https://doi.org/10.1007/s00737-026-01701-6.

Pudasainee-Kapri, S., Shrestha, T., Dahan, T., Wunnenberg, M., 2024. Translation, validation, and factor structure of the nepali version of Postpartum Bonding Questionnaires (PBQ-N) among postpartum women in Nepal. PLOS Glob. Public Health 4 (7), e0003469. https://doi.org/10.1371/journal.pgph.0003469.

Reise, S.P., 2012. The rediscovery of bifactor measurement models. Multivariate Behav. Res. 47 (5), 667–696. https://doi.org/10.1080/00273171.2012.715555.

Rodriguez, A., Reise, S.P., Haviland, M.G., 2016. Evaluating bifactor models: calculating and interpreting statistical indices. Psychol. Methods 21 (2), 137–150. https://doi.org/10.1037/met0000045.

Saur, A.M., Sinval, J., Del-Ben, C.M., Batista, R.F., Da Silva, A.A., Barbieri, M.A., Bettiol, H., 2023. The Postpartum Bonding Questionnaire: validity evidence from the Brazilian version. J. Child Fam. Stud. 32 (6), 1776–1788. https://doi.org/10.1007/s10826-022-02406-x.

Streiner, D.L., Norman, G.R., Cairney, J., 2015. Health Measurement Scales: A practical Guide to Their Development and Use, 5th ed. Oxford University Press.

Suetsugu, Y., Honjo, S., Ikeda, M., Kamibeppu, K., 2015. The Japanese version of the Postpartum Bonding Questionnaire: examination of the reliability, validity, and scale structure. J. Psychosom. Res. 79 (1), 55–61. https://doi.org/10.1016/j.jpsychores.2015.02.008.

Tabachnick, B.G., Fidell, L.S., 2019. Using Multivariate Statistics, 7 ed. Pearson.

Van de Schoot, R., Lugtig, P., Hox, J., 2012. A checklist for testing measurement invariance. Eur. J. Dev. Psychol. 9 (4), 486–492. https://doi.org/10.1080/17405629.2012.686740.

Vengadavaradan, A., Bharadwaj, B., Sathynarayanan, G., Durairaj, J., Rajaa, S., 2019. Translation, validation and factor structure of the Tamil version of the Postpartum Bonding Questionnaire (PBQ-T). Asian J. Psychiatr. 40, 62–67. https://doi.org/10.1016/j.ajp.2019.01.018.

Wittkowski, A., Vatter, S., Muhinyi, A., Garrett, C., Henderson, M., 2020. Measuring bonding or attachment in the parent-infant-relationship: a systematic review of parent-report assessment measures, their psychometric properties and clinical utility. Clin. Psychol. Rev. 82, 101906. https://doi.org/10.1016/j.cpr.2020.101906.

Wittkowski, A., Williams, J., Wieck, A., 2010. An examination of the psychometric properties and factor structure of the Postpartum Bonding Questionnaire in a clinical inpatient sample. Br. J. Clin. Psychol. 49, 163–172. https://doi.org/10.1348/014466509x445589.

Zijlmans, E.A.O., Tijmstra, J., van der Ark, L.A., Sijtsma, K., 2018. Item-score reliability in empirical-data sets and its relationship with other item indices. Educ. Psychol. Meas. 78 (6), 998–1020. https://doi.org/10.1177/0013164417728358.