Infant Behavior and Development 71 (2023) 101825

Contents lists available at ScienceDirect

Infant Behavior and Development

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

Review

Instruments to measure interaction of mothers and newborns: A systematic review

Tayomara Ferreira Nascimento [a,b,*], Silvia Cristina Mangini Bocchi [a], Rosa M. Trenado [b], M. Angeles Cerezo [b], Rodrigo Jensen [a]

[a] São Paulo State University (Unesp), Medical School, Botucatu, Brazil
[b] University of Valencia, Valencia, Spain

ARTICLE INFO

Keywords:
Neonate
Mother-Child Relations
Psychological Tests
Validation Studies
Behavioral Observation

ABSTRACT

Interactions between newborns and their parents/primary caregivers are characterized by asymmetric and dependent relationships. This systematic review mapped, identified, and described the psychometric parameters, categories, and items of instruments used to assess mother–newborn interaction. Seven electronic databases were accessed in this study. Furthermore, this research included neonatal interaction studies describing instruments' items, domains, and psychometric properties while excluding studies that focused on maternal interactions and lacked items for assessing newborns. Additionally, studies validated with older infants that did not have a newborn in the sample were used for test validation, which is a criterion used to decrease the risk of bias. Fourteen observational instruments from 1047 identified citations were included that addressed interactions using varying techniques, constructs, and settings. Particularly, we focused on observational settings that assessed interactions with communication-based constructs in the context of proximity or distance as influenced by physical, behavioral, or procedural barriers. These tools are also used to predict risk behaviors in a psychological context, mitigate feeding difficulties, and conduct neurobehavioral assessments of mother–newborn interactions. The elicited imitation was also an observational setting. This study found that the most described properties in the included citations were inter-rater reliability followed by criterion validity. However, only two instruments reported content, construct, and criterion validity, as well as a description of an internal consistency assessment and inter-rater reliability. Finally, the synthesis of the instruments reported in this study can guide clinicians and researchers in selecting the most appropriate one for their own application.

Babies (including full-term ones) physiologically detach from the maternal womb and are physically incapable and immature as they cannot perform effective behaviors to ensure their own survival. Furthermore, the first few days of the neonatal period are considered highly vulnerable as babies experience many physiological adjustments in their extrauterine life (Morton & Brodsky, 2016).

In the context of infant disability, the psychic evolution of a child is determined by impotence and initial helplessness. As such, newborns establish a psychological and biological symbiosis with their mother, particularly with the maternal breast, her body, and

* Correspondence to: Botucatu Medical School (Faculdade de Medicina de Botucatu), s/n, Av. Prof. Montenegro - Distrito de Rubião Junior, Botucatu, SP 18618-687, Brazil.
E-mail address: fenasta@alumni.uv.es (T.F. Nascimento).

https://doi.org/10.1016/j.infbeh.2023.101825
Received 16 September 2022; Received in revised form 9 February 2023; Accepted 14 February 2023
Available online 28 February 2023

the maternal psyche (Ainsworth, 1969). Considering this, the interactions between babies and their parents are marked by a naturally asymmetric and dependent relationship; that is, a relationship between a child who needs to be raised and an adult using certain socialization practices (Alvarenga et al., 2021; Cerezo et al., 2021). Neonates' contact with their parents or primary caregivers has both immediate and future consequences. For example, one study reported on the immediate benefits of skin-to-skin contact with the newborn for the overall family institution, the newborn themselves, the mothers, and the dyadic relationships involved therein (Zirpoli et al., 2019). These benefits include reduced hospitalization time, better thermal regulation, reduced pain during various procedures, strengthened bonding, improved sleep—especially deep sleep, which is considered a favorable strategy for neurobehavioral development (Zirpoli et al., 2019)—as well as improved parental identity and mood (Feldman et al., 2002; Ionio et al., 2021).

As such, current evidence shows that the interactions that occur during the prenatal period and the first few weeks of life have an undeniable impact on health. For example, chronic stress has been found to deteriorate neurons, consequently affecting the rest of the body (National Scientific Council on the Developing Child, 2020). Studies have also found that excessive and persistent adversity during early life can overwhelm biological systems, thereby causing long-term consequences with a critical impact on the lifestyle choices made in adulthood. Existing science has clarified two points in this context: a) occurrences during this period have both short- and long-term impacts on learning, behaviors, as well as physical and mental health; b) the interdependence of these domains is inexorably linked to the quality of the individual's physical and mental health (National Scientific Council on the Developing Child, 2020).

Research also shows that creating stable and responsive environments for children during their early years can prevent or reverse various negative conditions that would otherwise have consequences for their learning, behavior, and health (Eickmann et al., 2016). Therefore, it is important to consider that the caregiver ought to provide everything necessary for the baby to establish an age-appropriate sleep pattern and ensure that the infant receives sufficient attention and comfort when they wake up distressed. In addition, adults are the ones who establish conditions that can then promote well-being and hygiene, positive stimuli, and sources of pleasure (Trenado, 2001).

From birth, the mother engages in various behaviors that help her bond with her child. These processes are commonly described in the literature as vocalization, carrying and positioning the child in one's arms, visual and auditory stimulation, responding to the baby's crying, smelling, affection, and sensitivity, as well as developing adaptability to both the baby's needs and signals (Barnard et al., 1989; Norholt, 2020; White-Traut et al., 1988).

Thus, parents' perceptions and infant communication skills have been identified as elementary factors for establishing quality mother–infant interactions and are a primary indicator of child emotional regulation (Ainsworth & Bowlby, 1991; Bowlby, 1982). Through attuned, sensitive, and responsive interactions, parents or other primary caregivers can support and correct their baby's physiological, motor, and emotional arousal and activation (Alvarenga et al., 2021; Høifødt et al., 2020). Therefore, within these contexts, babies develop adequate social competence that guarantees them good relationships with others if they experience appropriate interactions (Cerezo, 1995; Rattaz et al., 2022).

Interaction observation instruments have been developed based on numerous theoretical assumptions and a series of constructs, such as responsive care, sensitivity, contingency, responsibility, emotional availability, and eye-to-eye contact, which research suggests is indicative of the quality of the relationship between the mother and child (Heimann, 1989). The abovementioned constructs can be used to evaluate the interactions of both reductionist (which evaluates each member of the dyad separately) and integral (which evaluates reciprocal exchanges between members) theoretical perspectives (Hernández et al., 2013; Horowitz et al., 2005).

The Ainsworth Maternal Sensitivity Scale (AMSS) is the gold standard for assessing parental sensitivity (Ainsworth et al., 1974). A review by (Mesman & Emmen, 2013) found 50 instruments conforming to Ainsworth's legacy. The AMSS is used to measure a mother's sensitivity to children aged three to 24 months (or thereabouts). Particularly, this instrument measures the mother's ability to accurately perceive their child's signals and to respond to those signals promptly and appropriately. Thus, observational coding is important and widely employed to assess neonatal interactions as it allows information to be obtained from the neonate and the caregiver, independently or simultaneously. Specific behavioral categories are central to observational measures in neonates because their repertoire of behaviors is limited in contrast to that of adults. Furthermore, the neonate's behavioral repertoire is related to their ability to react to and receive stimuli, to self-organize behavior, as well as to regulate sleep and wakefulness (Brazelton, 1973). Hence, evaluating the interactions of the neonate is a complex process.

In this regard, extant literature presents several constructs that examine interactive responses through paradigms such as the "Still Face" (Tronick et al., 1980) and the "Strange Situation." The latter evaluates children's responses to separation and union according to the attachment theory (Ainsworth, 1979; Bowlby, 1969, 1982).

However, in this research context, clinicians and other researchers often experience difficulty when choosing a tool. Rosenberg et al. (1986) suggest that measures to assess infant interactions should be reliable, even for behavioral patterns that seem ambiguous, such as those among young babies or those with disabilities. In addition, a system is required that is efficient, and that can easily be incorporated into a given intervention to help identify strategies that promote effective interactions between mothers/fathers and children.

Considering the wide variety of measurement instruments in healthcare, the Joanna Briggs Institute (JBI) has developed a set of theories, methods, and rigorous procedures to assess and synthesize different types of evidence (Hou et al., 2017). Studies based on the approaches proposed by the JBI have sought to obtain the best evidence available in this context. Specifically, these approaches have been developed through a complete and objective synthesis within each published scientific study, as well as from several relevant works using rigorous and transparent methods (Santos et al., 2018) .

For example, a panel of experts employed a consensus approach to develop the taxonomy, terminology, and properties of a measuring instrument called the "Consensus-Based Standards for the Selection of Health Measurement Instruments" (COSMIN) to

assist with the instrument analysis and selection process (Mokkink et al., 2010). Despite the existence of other methods, the COSMIN is a critical reference assessment tool in this field (Rosenkoetter & Tate, 2018) that can help assess the efficacy of instruments used for measuring the interactions between mothers and newborns.

As such, the current systematic review mapped, identified, and described the psychometric parameters, categories, and items of instruments that assess mother–newborn interactions.

1. Materials and methods

1.1. Study design

The present study conducted a systematic review to assess mother–newborn interaction instruments. It was designed based on the JBI guidelines for Systematic Property Measurement Reviews, the COSMIN, as well as on the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA; Page et al., 2021).

1.2. Eligibility criteria

This review focused on articles that describe or psychometrically assess instruments related to neonatal interactions during the neonatal period (which denotes the first weeks of the newborn's life), with no limits placed on the publication date. The languages for the studies were restricted to those within the author's domain, namely, Spanish, English, and Portuguese. Furthermore, this review included studies that 1) aimed to develop a new instrument, 2) used an instrument to measure study results, and 3) used an adapted version of a previously developed instrument. In addition, we only included articles that described the given instrument with sufficient characteristics with which to understand the aspects corresponding to the contribution of neonates to the interactions with their mothers or primary adult caregivers based on age-appropriate levels of physical and motor neurodevelopment.

Studies were excluded if the full text describing the items and domains of the instrument was not available. Furthermore, studies that proposed and/or were validated for another age group that did not assess age-appropriate interactive conditions in neonates (e.g., assessing vocalization and motor responses that are used to avoid an intrusive attitude) were excluded. Moreover, research investigating newborns with certain neonatal pathologies and studies that focused only on breastfeeding were excluded.

1.3. Study search and inclusion strategy

The first and last searches were conducted on November 10, 2021, and December 13, 2021, respectively. Subsequently, an update was performed on November 11, 2022. Furthermore, a psychology librarian was consulted to refine the search strategies. Based on the COSMIN guidelines for systematic reviews of patient-reported outcome measures (PROMs), the "Objectives" section of the PRISMA-P 2015 checklist was adapted by replacing "Participants, Interventions, Comparators, and Outcomes" with "Construction (neonatal interactions/mother/child relationships), Population(s) (newborns with their mothers in the hospital), and Instrument(s) and measurement properties (psychometric tests)."

The following electronic databases were used for the searches: PsycINFO, APA PsycTests, Scopus, PSICODOC, Cumulative Index to Nursing and Allied Health Literature (CINAHL), Embase, ProQuest, Medline, PubMed Central (PMC), and Web of Science. Excluding Pubmed, a search strategy was employed using broad keywords (e.g., "newborn," "baby," "mothers," "mother," "parents," "family," "scale," "questionnaire," "inventory," and "instrument"). The mesh terms were used with a measurement properties filter (a tool provided by the COSMIN, https://www.cosmin.nl/tools/pubmed-search-filters) in the electronic Pubmed database (Terwee et al., 2009). These search filters improved the completeness, efficiency, and quality of the literature search in this study. Furthermore, a filter was used in Scopus, CINAHL, PSICODOC, and APA PsycTests (age: 1 month). The full search strategy is available upon request.

Potential records that corresponded with the inclusion criteria were identified based on their title, acronyms, and names of the instruments mentioned. Thereafter, further literature searches were performed to find all original articles describing the respective measurement properties. Moreover, the reference lists of the included articles were analyzed.

Finally, consistent with our instrument search, two studies that conducted systematic reviews evaluating interactions between mothers, fathers, and newborns were reviewed. These studies were reviewed for instruments applicable to neonates that were designed with methods based on rating scales (17 tools; Munson & Odom, 1996) and observation instruments (23 instruments; Lotzin et al., 2015).

1.4. Study selection

The search results were imported into the Mendeley Desktop application (version 1.19.8; Elsevier; 2008; Glyph & Cog; USA) and subsequently into Rayyan to identify and remove duplicates. The results were also imported into (Ouzzani et al., 2016) to initiate peer reviews. Two reviewers with expertise in neonatal and family interactions made the decisions and resolved any differences through a consensus meeting with the other researchers. Furthermore, the reviewers specified the reason for the exclusion of any study.

A total of 1047 records were identified from databases based on the process shown in Fig. 1. Thereafter, 122 duplicates were removed from seven databases in Mendeley (n = 925). Subsequently, the references were imported into Rayyan. In the application, records were selected for inclusion in this review (n = 912) after running the duplicate check again. In the first review, when using the titles and abstracts, the reviewers included 71 studies, while they were undecided on others (n = 35) and excluded 806 records.

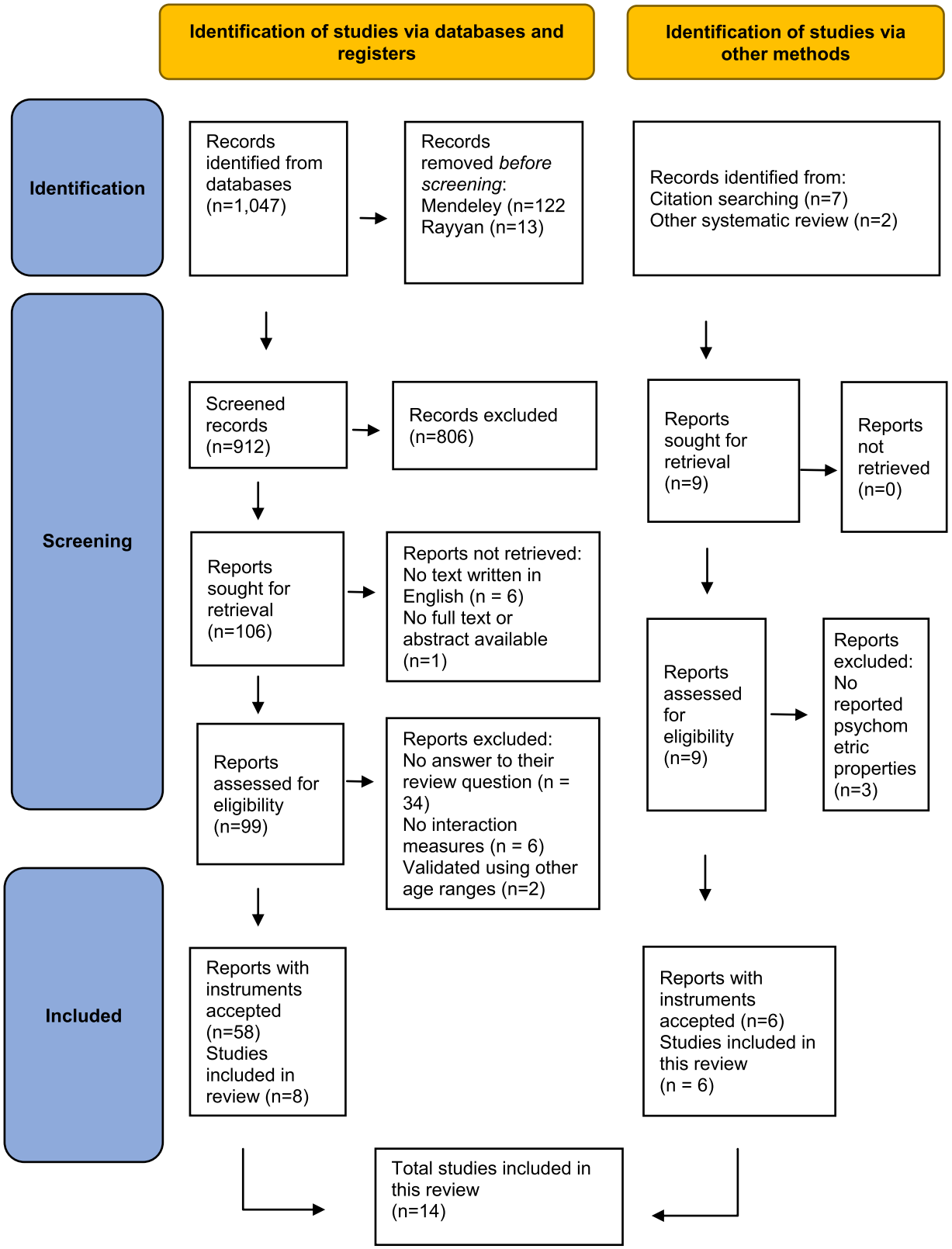

Fig. 1. PRISMA Flow Diagram, *Note.* Abbreviation: PRISMA = Preferred Reporting Items for Systematic Reviews and Meta-Analyses.

Following this, these records were reimported into Rayyan (n = 106) and assessed for eligibility using their full text by two independent reviewers. Additional records were included after conducting manual searches (n = 9). In total, 67 records obtained using the database (n = 58) and manual (n = 9) search strategies were evaluated for eligibility. From an analysis of 21 different instruments employed in these studies, a new literature search was performed using their names or acronyms in order to find each instrument's manual or the article with the description of their psychometric parameters.

Among these 21 instruments, seven were excluded as they outlined psychometric properties of instruments that studied infants older than 29 days or were validated with no newborns in the sample.

Regarding two other articles, the original studies could not be found. While an attempt was made to contact the authors, we had no response. Consequently, the authors decided to retain these studies in the review. In one study, the data were derived from a cultural adaptation and a validation study that was published in another language, whereas the other study did not describe the psychometric properties of the instrument. Thus, we obtained 14 studies that fulfilled the objective of this research.

1.5. Data collection and extraction

The reviewers extracted data from each study using a data collection form, which included the studies' data, design, purpose, sample population, instrument characteristics, and measurement properties. Additionally, the reviewers manually collected study characteristics from each report and organized them within a spreadsheet.

Furthermore, the reviewers applied the risk of bias from the COSMIN Checklist (Mokkink et al., 2010), which has been adapted to evaluate information about the measurement properties and methodological quality of each study. However, these are not described in this article.

1.6. Evidence synthesis

The results were analyzed qualitatively, with a descriptive framework being constructed. Thereafter, the studies were categorized according to their observation setting. Finally, a narrative description of the psychometric parameters was constructed. Data on validity and test accuracy were described according to the information presented in the respective manuals and publications.

For the psychometric parameters, five aspects were obtained regarding the information on the development and validation of the instruments described, including their content validation, construct validation, criterion validation, internal consistency, and inter-observer reliability.

This study considered validity measures that were related to the tests and concepts used to measure the assessment (Polit & Beck, 2019). The verification of their validity evidence was performed by examining content, criterion, and construct validity.

Content validity describes the degree to which an instrument has an appropriate sample of items for the construct being measured and whether the items adequately represent its content domain (Pasquali, 2001; Polit & Beck, 2019). Furthermore, content validity determines whether the choice of items is appropriate and relevant. Pasquali (2001) postulates that the criterion validity of a test refers to its degree of effectiveness in predicting the studied subject's performance. In contrast, construct validity has been understood as the "degree to which the test measures a theoretical construct or trait for which it was designed" (Polit & Beck, 2019).

Reliability, according to internal consistency, indicates the degree to which the various components of a multi-component instrument consistently measure the same attribute. Internal consistency is a widely recorded aspect of reliability and is estimated using the alpha coefficient (or Cronbach's alpha) index. For all these reliability indices, values closer to 1.00 suggest stronger reliability (Polit & Beck, 2019).

Inter-rater reliability is used to assess the degree to which two different raters or observers obtain the same score when measuring an attribute. When the ratings are dichotomous (referring to the presence or absence of a specific attribute), the preferred index is Cohen's Kappa coefficient, which also has values ranging from 0.00 to 1.00 (Polit & Beck, 2019).

2. Results

2.1. Characteristics of instruments on neonatal interactions

A total of 14 articles were included to answer this study's research question. The instruments cited in these studies addressed neonatal interactions through different techniques, constructs, and settings. The following paragraphs describe the relevant features of the instruments developed to observe interactions between newborns and mothers or with any other adult primary caregiver. Furthermore, the data synthesis from the psychometric evaluations conducted for each instrument's construction is underscored.

The current review found that all reports indicated the need for training and education. Furthermore, five (33%) instruments mentioned the existence of training offerings and official courses with certificate issuance and paid access (Anderson et al., 2004; Baird et al., 1992; Costas Moragas et al., 2007; Feldman, 1998; Provenzi et al., 2018).

The observational settings and psychometric parameters of the instruments are presented below according to the data reported. The reviewers grouped the instruments into two mutually exclusive categories based on the scenario employed for their validation that were distributed at the researchers' discretion. The first category comprises instruments related to interactional and communicative and neurobehavioral acts. The second category lists instruments that address the issues concerning contact, physical barriers, and eating difficulties.

2.2. Narrative synthesis of the interactions: communicative and neurobehavioral acts

A total of nine instruments investigated newborn interactive acts in their assessment instruments (Baird et al., 1992; Bakerman & Brown, 1977; Brazelton, 1973; Censullo et al., 1987; Heimann, 1989; Kumar & Hipwell, 1996; Provenzi et al., 2018; Raack, 1989; Rime et al., 2018). The following section reports the observational settings and psychometric parameters of these instruments.

2.2.1. Neonatal behavioral assessment scale

Brazelton (1973)—the author of the Neonatal Behavioral Assessment Scale (NBAS), Newborn Behavioral Observations (NBO), and co-author of the Mini Dyadic Code scale (DMC; Censullo et al., 1987)—states that the newborn is competent, active, and capable of responding to surrounding stimuli. Particularly, newborns possess superior neurological centers that serve to modulate their reactions. Herein, NBAS is widely used to assess neonatal neurobehavioral functioning. Thus, Brazelton (1973) developed a technique based on interactive evaluations that is predictive of responses to infants' behavior in each situation.

The content validity was not reported even though a validity construct was found. The authors' mention formed one of the main qualities attributed to this scale. Criterion validity was demonstrated within the observing motor performance scale that was significantly correlated with the Prechtl total Score (Pearson correlation coefficient $r = -0.41$). However, poor correlations were found with the habituation and state items of the NBAS (Majnemer, 2008). Regarding reliability, internal consistency ($r = -0.15$ to 0.32), test-retest ($r = -11$ to 0.52), and inter-rater reliability (60–100% agreement) were all reported (Majnemer, 2008).

The NBAS reliability data obtained through a validation study for the Spanish language reported a coefficient between moderate and high. Particularly, the average Cronbach's alpha coefficient was 0.78 (Costas Moragas et al., 2007).

2.2.2. Behavioral dialogues assessment instrument

According to the authors of this instrument, a baby's communicative acts refer to behaviors that may increase responsiveness to and from the mother (Bakerman & Brown, 1977). Therefore, the Behavioral Dialogues Assessment Instrument, which is an observational coding instrument, analyzes patterns of dialogical interactive behaviors wherein the two parties contribute. Bakerman and Brown (1977) developed an instrument that views mother–newborn interactions as a dialogue in "communicative acts that do not require communication" and described them using simple and transient probabilities.

However, content validity was not reported. The authors did report that their constructs and criterion validity discriminated newborns and mothers in an interpretable, sensitive, and consistent manner compared to other results from at least two other researchers using similar variables. Furthermore, the researchers only cited the authors of the instruments that were not cited in this review as the study was not available. Regarding reliability, internal consistency was not reported. Moreover, the minimum inter-observer reliability stipulated for the study had a pre-determined criterion of 75% agreement over three successive observations.

2.2.3. Dyadic mini code

The DMC was developed based on various important components of synchronous interactions. Particularly, the attachment theory provides a theoretical basis for the DMC (Ainsworth et al., 1978; Bowlby, 1969). In this instrument, dyadic or low synchrony is obtained through observations of the reciprocal dyadic interactions regarding the articulation of behaviors, and in a manner that is not separated from the individuals, as in other instruments.

Content and construct validity, however, were not reported. Criterion validity was demonstrated by observing that the instrument discriminated preterm and term dyads according to the expected differences. However, the authors reported that further refinement of the DMC is needed (Censullo et al., 1987). Its internal consistency was not reported, while the inter-rater reliability was determined by a Cohen's Kappa and percentage agreement. Specifically, the Kappa was 0.86 for the total score and ranged from 0.63 to 0.92 for each of the six items (Censullo et al., 1987).

2.2.4. Neonatal imitation, gaze aversion, and mother–newborn interactions

This assessment instrument examines the relationship between elicited imitations in newborns and their social interactions. Furthermore, this particular assessment studies neonatal imitation, gaze aversion (GA), and mother–newborn interactions. Elicited imitation describes a language sampling procedure in which a child is asked to repeat a statement modeled by the examiner (Heimann, 1989).

However, content and construct validity were not reported. Regarding reliability, its internal consistency was also not reported. Reliability was calculated using Cohen's Kappa. In contrast, criterion validity was reported as correlations found in children's behavior

during their facial actions at the age of two to three days, with imitation scores being observed at three months. Their tongue protrusion, mouth opening, and state coding yielded coefficients of 0.87 ($p < 0.001$), 0.77 ($p < 0.001$), and 0.79 ($p < 0.001$), respectively, with these, then comprising the items evaluated in the instrument (Heimann, 1989).

2.2.5. Mother-infant communication screening

The Mother-Infant Communication Screening (MICS) was developed to assess communicative interactions between newborns and mothers. Raack (1989), a phono-audiology specialist, described connections between maternal language and fetal movement that continue throughout infancy. Both verbal and non-verbal components of language readiness and their relationship to maternal interactions underscore the MICS's items. The authors found that the instrument is sensitive to the early detection of inappropriate interactive patterns that could compromise children's development (Galván-Bovaira et al., 2002).

Construct validity and content validity were, however, not reported. Criterion validity was calculated against the Nursing Clinical Assessment Tool (NCAT) as an external measure. Consequently, a positive and significant correlation ($r = 0.504$, $p < 0.001$) was found for the total score, as well as one for the two most comparable subscales ($r = 0.492$, $p < 0.001$). Internal consistencies of 0.89–0.94 (MICS total score) were reported depending on the choice of subscales used, while 0.79 was found for the NCAT (Byrne & Keefe, 2003). Inter-rater reliability was found to be 0.85.

2.2.6. Infant-parent social interaction code

The Infant-Parent Social Interaction Code (IPSIC) was designed to provide a reliable and useful profile of the interaction process between parents and babies (Baird et al., 1992). Accordingly, the IPSIC detects contingency behaviors, intrusions, and communicative acts, as highlighted by specialists, that then comprise an instrument that aims to facilitate the articulation between evaluations and interventions. Furthermore, the authors describe the limitations of certain instrument constructs in terms of reliability measures. For example, the production of an artificially inflated Kappa can be due to the relative frequency of behaviors and the agreement profile. In addition, they refer to authors who underscore the usefulness and applicability of the IPSIC as a tool. However, as with any reasonable instrument, the IPSIC should be used with caution.

Construct validity was reported, while content and criterion validity were not. Furthermore, internal consistency was not reported. Reliability was calculated using Cohen's Kappa. The findings indicate sufficient internal consistency with Cronbach's alpha values of 0.90 (range=85–97) and Cohen's low Kappa values of 0.47 (range=27–68; Baird et al., 1992).

2.2.7. Bethlem mother-infant interaction scale

The Bethlem Mother-Infant Interaction Scale was developed in a psychiatric context (Kumar & Hipwell, 1996). This scale aims to predict maternal psychological changes that may pose a danger to the baby (e.g., possession) in order to provide prevention.

Construct, content, and criterion validity were all reported. Furthermore, the instrument reported good validity criteria and high internal consistency with a Cronbach's alpha of 0.93, as opposed to other methods used in non-clinical settings. The Kappa values reported ranged from 0.52 (eye contact) to 0.76 (vocal contact).

2.2.8. Family alliance assessment scales for diaper change play

This scale was designed to assess the quality of triadic family interactions and was adapted for application during the first week of the baby's life based on a perinatal triadic family assessment instrument, the "Prenatal Lausanne Trilogue Play," for use among two- to three-month-old infants (Favez, 2006, as cited in Rime et al., 2018). Its observations were conducted under the free-play paradigm that was conducted in a diaper-changing situation.

Construct, content, and criterion validity were all reported. The instrument had a sufficient internal consistency of 0.91 and a good consensus among the three raters (ICC=0.76; range=66–82), with the correlations all being significant ($p < 0.05$; Rime et al., 2018).

2.2.9. Neurobehavioral scale of the Network of Neonatal Intensive Care Units

The Neurobehavioral Scale of the Network of Neonatal Intensive Care Units (NNNS) assesses the neurological integrity of newborns by measuring neurobehavioral patterns (e.g., habituation and regulatory status) and other factors common to both instruments (Provenzi et al., 2018). The main difference between the two neurobehavioral assessment instruments mentioned in this review is that the NBAS provides an evaluation of "Smiles" (i.e., the number of times the newborn expresses a smile), whereas the NNNS comprises a Stress/Abstinence scale.

Construct and content validity were not reported. However, its criterion validity showcased that it could assess neurobehavioral performance in healthy infants. The authors also found good internal and concurrent validity. This instrument demonstrates sufficient internal consistency and concurrent validity without presenting quantitative data. Thus, two instruments were found for observing neonatal behaviors with a focus on evaluating the interactions and documentation of childhood behaviors, neurological integrity, and neurobehavioral functioning, specifically, the NBAS (Brazelton, 1973) and the NNNS (*Provenzi et al., 2018).

2.3. Narrative synthesis of interactions: contact, physical barriers, and feeding difficulties

Five instruments were described in this study that fit into its second category (Anderson et al., 2004; Dumas et al., 2013; Feldman, 1998; Santos, 2008, 2010).

2.3.1. Coding interactive behavior

The Coding Interactive Behavior (CIB) is a behavioral coding instrument with several versions assessing behaviors in different groups (e.g., between couples). In the neonatal version (Feldman, 1998; Silberstein et al., 2009), the CIB codes behaviors by observing the interactions of dyads in a play situation or in a more elaborate context. The instrument was developed with low-risk premature infants to predict the development of feeding difficulties during their first year of life.

Content and construct validity were not reported, while a criterion validity was found that indicates this instrument is sensitive when assessing the interactions between parents and low-risk premature infants, as well as in predicting the development of feeding difficulties during the first year of life. Thus, these findings reveal that negative infant relationship patterns can predict future feeding relationships in their first year. Furthermore, the rhythm of maternal behaviors can be analyzed as an adaptive factor to enhance a smooth transition to oral feeding. Internal consistency, however, was not reported. Interobserver reliability was determined with a Cohen's Kappa of 0.84 (range=78–91) for situations wherein the child was being fed and 0.84 (range=81–87) in non-feeding situations (Silberstein et al., 2009).

2.3.2. Index of mother-infant separation

Through observations, the Index of Mother-Infant Separation (IMIS) codifies and documents the nature of mother–infant contact and postpartum separation via measuring contact versus separation while also discriminating between groups as well as selecting contact behaviors that facilitate mutual care. Initially referred to as the MICS, the instrument was transformed into the IMIS upon discovering that two other scoring systems with the same name were in use (Anderson et al., 2004).

The IMIS content validity scores for each item determined by perinatal experts ranged from 77% to 100%. Structural validity was supported using known-group hypothesis testing ($p < 0.0001$). Furthermore, criterion validity was also reported. No report on its internal consistency was found, whereas extant research reported an inter-rater Kappa reliability of 0.86–0.90. (Anderson et al., 2004).

2.3.3. The scale of observation of the mother-baby bond interned in the NICU

The Scale of Observation of the Mother-Baby Bond has been interned in the neonatal intensive care unit (NICU). This scale was developed to measure parental stress related to the physical and psychosocial environment measured within the NICU.

Content and construct validity were not reported. However, criterion validity for the instrument was calculated. Particularly, sensitivity (0.75), specificity (1), positive predictive value (1), negative predictive value (0.84), and total correct predictions (0.89) were all reported. Furthermore, internal consistency was reported. Additionally, a high correlation was observed in the dyadic total scores between both observers for inter-observer reliability (rptotal=0.99 $p < 0.001$); likewise, this was found in the total scores of the mother and the baby (rpmmother=0.97 $p < 0.001$; rho Spearman's pbaby=0.99 $p < 0.001$; Spearman's rhobaby=0.99 $p < 0.001$). In the dyadic functional scores, high correlation indexes were observed with values ranging between 0.97 and 1.0 with a statistically significant level of $p < 0.001$, except for the Postural function (rF. Postural=0.70 $p > 0.05$) (Santos, 2008).

The same author published a scale comprising similar item descriptions. Research suggests that another scale—the "New observation scale of the bond mother-baby incubator" (N-EOV-INC)—is a later version of this mentioned scale (Santos, 2010). The N-EOV-INC includes one more domain than does that mentioned in Santos (2008) called "Contact with the incubator." In relation to the N-EOV-INC, this review found no studies containing descriptions of this instrument's psychometric data, even after repeated attempts to contact the authors. However, the instrument was retained in this study based on the limited instruments available and minor variations between the version developed by the same author (Santos, 2008).

2.3.4. Observational assessment of mother-baby interaction

This instrument can be used to evaluate the interactions between newborns and mothers. However, contact herein was limited by that conducted with blankets and excessive clothing (*Dumas et al., 2013).

Content and construct validity were both reported. The authors also found cross-cultural validation and inter-observer reliability by expert panels for facial and content validity. Furthermore, relevant findings demonstrated criterion validity wherein the predictive value was analyzed by comparing groups of babies in traditional Russian attire (swaddled). These findings demonstrate that sensitivity was assessed in terms of differences in kindness and rudeness according to the occurrence of each kind of behavior. Internal consistency was not reported. However, inter-observer reliability of at least 0.80 has been reported for each item.

Table 1 shows the descriptions of the domains and items of the instruments and other relevant information.

Table 1
Instruments for Observing Behavior and Mother–Newborn Interactions.

Instrument	Measure Age Purpose/Focus	Domain Section	Sample Original Population Recall Period Number of Items
Brazelton (1973) NBAS Costas Moragas et al. (2007)	Scoring System Age: Not Reported Identify the full range of individual neurobehavioral functioning and identify areas of difficulty.	(1) Autonomic and (2) Motor Systems, (3) Habituation, (4) State Organization, (5) State Regulation, and (6) Social Interactive. (7) Complementary items intended to describe the most qualitative aspects of the newborn's performance during the examination and (8) Smiles (number of times the neonate sketches this expression). All behavioral items (except Smiles) are scored on a 9-point scale, with 9 being the most ideal. However, 8 items are scored according to a curvilinear scale, wherein the central scores are optimal. These are recoded according to a 5, 6, or 8-point linear scale. The scores are specific to each item. Reflex responses are scored on a 4-point scale. With this scale, a single score is not obtained, but a profile of scores are calculated that describe the neurobehavioral organization of the newborn.	220 newborns Spain Healthy full-term neonates No report period 35 behavioral items
Bakeman and Brown (1977) (Not reported)	Observational coding system Protocol Age: Not Reported To describe the interactive patterns of mothers and their newborns.	Domains: (1) State: mother and infant concurrent engaging in communicative acts; (2) The mother-alone state: mother acting alone; (3) The infant-alone state: infant acting alone, and (4) The quiescent state: neither acting. The infant's communicative acts: domains: (1) SWIP-infant swipes at his mouth or sucks his fingers; (2) RFUS-infant refuses the bottle; (3) DRIB—milk dribbles out of the infant's mouth; (4) TRAIN infant trembles or roots in the absence of nipple stimulation; and (5) IVOC-infant vocalizes, typically by making sucking noises or by whimpering. Infant characteristics were rated with a 9-point scale on the following items, which were selected from the Graham-Rosenblith scale: (1) motor strength, (2) tactile adaptivity, (3) visual responsiveness, and (4) auditory responsiveness.	45 dyads USA The behaviors of the mothers and the infants were recorded for 8 s on the third day of the infant's life, observed for two 1/2-hour sessions, one at noon and one at 4:00 pm, during which time the mother bottle-fed her infant in her hospital room. 100 hierarchically structured codes
Censullo et al. (1987) DMC	Scoring System Age: 0–6 months To measure levels of synchrony in early infant–adult interaction.	Items: (1) mutual attention, (2) positive affect, (3) mutual turn-taking, (4) maternal pauses, (5) infant clarity of cues, (6) maternal sensitive responsiveness of the infant. Each item is given a score of 1 or 2 and a total score rated as high or low synchronous. The total score ranges from 6 to 12. A score of 6–9 is ranked as low synchronous, with 10–12 ranked as synchronous.	20 dyads USA Term and preterm infants and their mothers in a hospital Recall period not reported 6 items with a total score summary rating trained applicators in paid courses

(*continued on next page*)

Table 1 (*continued*)

Instrument	Measure Age Purpose/Focus	Domain Section	Sample Original Population Recall Period Number of Items
Heimann (1989) (Not reported)	Observational coding system Age: Not reported To examine whether a relationship is elicited when studying neonatal imitation, gauze aversion, and mother–infant interaction.	Codes to imitative responses and state. (1) Imitation of tongue protrusion, (2) Mouth opening, and (3) Lip protrusion of baby. The study also uses another instrument named the Monadic Phases for three-month-old babies. 11 Categories were used: (1) High Protest, (2) Protest, (3) High Avert, (4) Avert, (5) Pick-me Up, (6) Object Attend, (7) Social Attend, (8) Object Play, (9) Social Play, (10) Talk, and (11) Positive Avert. Three codes were assessed in all three observations at two to three days of age, in the home at three weeks, and in a laboratory setting at three months of age.	32 full-term infants USA Videos of full-term infants (14 male and 18 female) Recall period not reported Three Codes and Monadic Phases
Raack (1989) MICS Galván-Bovaira et al. (2002) Byrne and Keefe (2003)	Scoring System Age: 0–12 months To assess communicative interaction between mothers and their children from 0 to 12 months.	(1) Language and Synchrony, (2) State of Discomfort, (3) Feeding, (4) Game or Neutral State, and (5) Rest. Subscale from 1 to 5 for each item (Total score/no. of items administered). Final score ranging from 1 to 5.	10 dyads Spain Mother/infant dyads Recall period not reported Five domains and a total score
Baird et al. (1992) IPSIC	Observational coding system Age: Not Reported To measure infant, parent, and dyadic aspects of interactions that are sensitive to recent empirical findings and could facilitate the link between interaction assessment and intervention.	(1) Parent response contingency, (2) Parent directiveness, (3) Parent intrusiveness, (4) Parent facilitation, (5) Infant initiation, (6) Infant participation, (7) Infant signal clarity, (8) Intentional Infant, and (9) Communicative acts. The first five minutes of the tape are considered a warm-up period and are not coded. Beginning at 10 min, each 15-second segment of the videotape was viewed from Minute 5 and continued to Minute 10. Thereafter, the tape stopped and the presence or absence of each behavioral construct was recorded. Segments are viewed as many times as the observer chooses. A " + " indicates the presence, while a " - " indicates the absence of behavioral constructs for specific observation intervals.	159 dyads EUA Infants ranging from birth to 31 months Mothers and infants were videotaped in either their homes or a home-like clinic setting. Parents were asked to "Play as you normally do when you are not feeding, bathing, or changing your infant," while a 10-minute sample of interactions was videotaped. Four parent variables, four infant variables, and one dyadic variable
Kumar and Hipwell (1996) BMIS	Scoring System Age: Not Reported To measure mother–infant adjustment in mother-baby units.	Domains: (1) Eye contact, (2) Physical contact, (3) Vocal contact, (4) Mother's mood: a) Dialogue score (Sum 1–4), (5) General routine, (6) Assessment of Risk: Total score (Sum 1–6), and (7) Baby's contributions to interaction. Of the seven domains: (1) Four different measure aspects of the mother's contribution to the dialogue with her baby, (2) One measures her capacity to organize and maintain routine care, (3) One attempts to rate the perception of the staff of risk to the child on which assessments were made of the mothers' general ability to manage a daily routine and her competence in caring for the baby's physical needs.	78 dyads England Mothers with a severe mental illness in the postpartum period were admitted to the psychiatric unit at Bethlem Royal Hospital, where five minutes of interaction between them and infants were filmed and coded during play or feeding Seven items in two subscales

(*continued on next page*)

Table 1 (*continued*)

Instrument Measure Age Purpose/Focus	Domain Section	Sample Original Population Recall Period Number of Items
	The ratings were made on a 5-point scale. A score of 0 indicates that the mother interacts with her child inappropriately, sensitively, and in a well-organized way, respectively. A score of 3 indicates severe disturbances demonstrating that most of the time that the mother is with her baby, she was unable to sustain any meaningful dialogue or interactions; and (4) the remaining subscale rates the baby's contribution to their interaction.	
Feldman (1998) CIB Silberstein et al. (2009) Observational coding system Age: Not Reported To analyze social interactions between two or more partners. The instrument was used in non-feeding play interactions.	The codes: (1) Mother's touch: affectionate (kissing, caressing, hugging, gently stroking, or mother touching infant with clear positive affect); (2) Functional (wiping mouth, arranging clothes or blanket), and (3) No touch (mother just holds the baby); (4) Mother's averts her gaze from infant to bottle; (5) Aversion gaze (no eye contact with infant or bottle); (6) Infant feeding performance: robust, weak, or not feeding. The system contains one infant and four mother codes. In addition, maternal adaptation was coded globally on a scale of 1 (low) to 5 (high).	76 dyads Location not reported Premature infants and their mothers were recruited during the first two weeks following birth in a tertiary care NICU. Videotaped during six minutes for feeding and seven minutes for non-feeding play interactions were 11 items
Anderson et al. (2004) IMIS Observational coding system Age: Not Reported To measure the process of mother–infant contact or separation post-birth.	Five behavioral groups defined the types of contact and separation: (1) Mother–infant contact, (2) Father–infant contact, (3) Staff nurse contact, (4) Other contact; (5) Nurse researcher contact (and separation and if rooming-in).	224 newborns USA Healthy newborn infants 1 h/Time-sampled observations occurred every 15 min by trained applicators in paid courses 37 items
Santos (2008) (Not reported) Scoring System Age: Not Reported To evaluate the interactions of behavioral behaviors between the mother and her baby.	Five functions: (1) Approach, (2) Corporal, (3) Visual, (4) Verbal, and (5) Posture function. Functional and total scores are obtained for the mother (a), the baby (b), and the dyad (a+b). Subsequently, the numbers of the marked items are added, obtaining the Total NEA (total encounter level).	43 dyads Argentina Mother–baby dyads in UCIN of three hospitals. It is measured from the moment the mother approaches the incubator with the baby inside and lasts until the end of the visit or until five minutes have elapsed Five items and a total score
Santos (2010) N-EOV-INC Scoring System Age: Not Reported To detect indicators of risk in the mother–baby bond during confinement in the NICU	This instrument contains infant 6 categories: (1) Approach, (2) Contact with the incubator, (3) Body, (4) Visual, (5) Verbal, and (6) Postural function.	106 dyads Argentina Mother–baby dyads Not available in summary 21 items Six categories
Dumas et al. (2013) (Not reported) Rating Scale Age: Not Reported To measure and explore maternal behaviors during breastfeeding Session 4 days after childbirth and the influence of perinatal care routines such as skin-to-skin contact, baby's apparel, and separation on mother–infant interaction post-birth.	Instrument to assess the state of neonatal wakefulness according to the behavioral state of Brazelton and Nugent (1995): (1) Deep sleep, (2) Light sleep, (3) Drowsy, (4) Alert, (5) Alert and active, (6) Crying. This evaluation is carried out in two moments. Adult assessment: (1) Mother's	151 dyads Russia During the breastfeeding period, at postpartum day Four in the mothers' rooms during 25–45 min 12 items and field notes

(*continued on next page*)

Table 1 (*continued*)

Instrument Measure Age Purpose/Focus	Domain Section	Sample Original Population Recall Period Number of Items
	movements to and with the baby; (2) Mother's attempts at latching; (3) Mother's voice; (4)Mother's patience; (5) Mother's stimulation (Types of mother's stimulation if yes); (6) Stimulation of the baby by staff (Types of stimulation by staff if yes, otherwise the mother models staff); (7) Mother's pain at latching; (8) Mother's general affective responsiveness to the baby; (9) Eye-to-eye contact; (10) Attempts at eye-to-eye contact; (11) Movements indicating avoidance of baby; (12) At least one effective breastfeeding sequence. The evaluator responds to each statement to qualify the predominant behavior observed. Specifically, seven items are assessed on a scale from 1 to 5 (from the roughest to the gentlest behavior); two items are on a 3-point scale, and three items are answered "yes or no" according to the occurrence or nonoccurrence of the behavior.	
Provenzi et al. (2018) NNNS Scoring System Age: Not Reported Assesses at-risk infants (particularly substance-exposed), documenting neurological integrity and the broad range of behavioral functioning.	(1) Habituation, (2) Attention, (3) Arousal, (4) Regulation, (5) Handling procedures, (6) Quality of movement, (7) Excitability, (8) Lethargy, (9) Nonoptimal reflexes, (10) Asymmetric reflexes, (11) Hypertonicity, (12) Hypotonicity, and (13) Stress/abstinence scale. 12 packages with 45 items (scores range 1–11) and seven stress scale packages with 70 items (yes/no scores).	99 full-term infants EUA Clinically healthy full-term infants were recruited from a well-child nursery. Recall period not reported 128 items and 13 summary scores
Rime et al. (2018) FAAS-DCP Scoring System Age: Not Reported To assess the quality of family relations observed during mother–father–infant triadic interactions	Nine interactive dimensions: (1) Readiness to interact, (2) Gaze orientation, (3) Inclusion of partners, (4) Co-parental coordination, (5) Role organization, (6) Parental scaffolding, (7) Shared and co-constructed activities, (8) Sensitivity, and (9) Family warmth. The nine interactive dimensions are rated on a 5-point scale, with a score of 5 representing optimal functioning and a score of 1 representing significant dysfunction.	44 triads Switzerland Triads (mother, father, and neonate) Recall period not reported Nine items

Note. United States of America **(**USA), Bethlem Mother-Infant Interaction Scale (BMIS), Coding Interactive Behavior (CIB), Mini Dyadic Code scale (DMC), Family Alliance Assessment Scales for Diaper Change Play (FAAS-DCP), Index of Mother-Infant Separation (IMIS), Mother/Infant Communication Screening (MICS), Mother-Infant Togetherness Survey (MITS), Neonatal Behavioral Assessment Scale (NBAS), Network of Neonatal Intensive Care Units (NNNS), New observation scale of the bond mother-baby incubator (N-EOV-INC), and Infant-Parent Social Interaction Code (IPSIC).

3. Discussion

The results from this review of the aforementioned studies indicate that most current instruments evaluate interactions in the context of communication, proximity, or distance while predicting risk in the bond between mothers and babies due to physical, behavioral, or procedural barriers. Instruments that observe interactions during breastfeeding were used to predict behaviors associated with psychological risk in a psychiatric care context and to predict eating difficulties within the course of one year. In addition, studies on the neurobehavioral aspects of infants and their ability to interact have been reported.

Except for the instrument by Santos (2010), measurement properties were included that represent essential aspects of measurement

quality (reliability and validity) of the mapped instruments, mostly described in terms of inter-observer reliability (13), followed by criterion validity (12). Furthermore, studies that assessed construct validity (7), internal consistency (5), and content validity were the least reported (4).

Among the five psychometric parameters assessing validity and reliability (construct, content, and criterion validity, as well as measurement statistics for internal consistency and reliability) were found in two studies, namely that on the Bethlem Mother-Infant Interaction Scale (Kumar & Hipwell, 1996) and the Family Alliance Assessment Scales for Diaper Change Play (Rime et al., 2018). The instruments that assessed four of these measures were: the NBAS (Brazelton, 1973), the IMIS (Anderson et al., 2004), and the Observational Assessment of Mother-Baby Interaction (Dumas et al., 2013). The NBAS (Brazelton, 1973) is an interactive assessment technique that is considered among the most suitable tools for detecting deficits and identifying emerging abilities in the neonate (Costas Moragas et al., 2007).

Overall, a good instrument in this context should describe psychometric conditions, which can serve as an indication of its validity and reliability. In contrast to commonly performed evidence evaluations (Terwee et al., 2007), this study has not summarized the criteria of quality scores into an overall finding. A quality score often assumes that all quality parameters are equally important; however, this is not always the case. Inter-observer reliability presents a crucial issue when employing observational analysis instruments used to assess the interactive aspects of newborns. Particularly, the observer challenge and the difficulty in achieving levels of inter-observer reliability increase when the baby is younger.

Although the instruments seem to be of good clinical use, the lack of evaluation descriptions performed by the authors of each study allows the inference of those psychometric evaluations in studies evaluating inadequately described neonatal interactions. The existing assessment procedures presented in this review assess neonatal interactions based on observational coding procedures and diagnostic classification systems. In addition to observing the psychometric parameters that facilitate the selection of a viable instrument for the required item of measure, tool selection will depend on the behavior being assessed. In addition, both the personnel available to use the instrument and the purpose of the instrument itself will guide the selection of the most appropriate tool in a clinical or research context.

The current review found instruments that assess the usefulness of interaction reports from before the baby is even one month old. However, several items were not applicable to neonates, such as the Parent-Infant Interaction Observation Scale (PIIOS), which was developed for infants aged two to seven months (Svanberg et al., 2013). Particularly, (Naughton et al., 2019) found that three out of 13 items were unlikely to apply to two-week-old infants while they were awake.

A challenge encountered by several authors was that an instrument could be used from age zero; however, they were only validated with older infants. Studies designed to assess the interactions of the typical newborn beyond the setting of the intensive care unit were limited. Particularly, settings characterized by other intervening factors, such as prematurity or diseases, can affect the assessment of interactions.

Reeb-Sutherland et al. (2022) underscored advances in understanding interactions between caregivers and children, using different methodologies, multi-methods with sensors (Guida et al., 2021), "wearable devices" (Sadeh & Acebo, 2002), "hyper scan" (Perone et al., 2020), and "cameras and eye trackers" (Lourenço et al., 2021). However, no tool has been found that combines innovative technologies with the use of various methods to aid in the observation of infant–caregiver interactions; therefore, this gap still needs to be filled.

Future research needs to study the interactions between mothers and babies during this critical phase during the first 27 days of life. Preferred studies should be associated with the use of technological resources such as "wearable devices," scanners, sensors, and data capture images that can facilitate the process of obtaining, recording, processing, and analyzing data for investigation and clinical use.

3.1. Limitations and future directions

This review only included studies of certain languages and excluded all others, meaning that some important works may have been missed. Furthermore, the psychometric data evaluation was not extended to other studies that used the aforementioned instruments.

The data obtained in this study indicate a need for developing a new instrument that evaluates the interactions between mothers and newborns while considering temporality in the dyadic context. Furthermore, knowledge underscoring the importance of observer training and psychometric parameters, as described by these 14 instruments, will inform the development and study of the psychometric parameters of the intended instrument. Particularly, this could then support the need to develop a learning protocol that ensures the adequate training of reliable observers in the future.

4. Conclusion

This systematic review summarized studies that address the issue associated with mother–newborn interactions. Instruments were identified, and a description of their usefulness and reported domains of psychometric parameters were provided.

The present study underscored 14 instruments that assess interactions and that can be used to predict risks in mother–newborn bonding. Particularly, these instruments consider the effect of physical barriers, such as incubators and clothing, while assessing several aspects of bonding and separation, communication, neurobehavioral aspects of newborns, and the psychological aspects of mothers, among others.

This study compiled and synthesized data that provide an overview of useful instruments for studies related to interactive behaviors with newborns and their families during the first weeks of a newborn's life.

Funding

This work was supported by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001, and by the University of Valencia. The funding sources had no involvement in the study decisions.

CRediT authorship contribution statement

Tayomara F. Nascimento: Data curation, Formal analysis, Writing – original draft, Visualization. **Silvia C. M. Bocchi:** Data curation, Writing – review & editing, Resources. **M. A. Cerezo:** Conceptualization, Methodology, Writing – review & editing, Formal analysis, Supervision. **Rosa M. Trenado:** Conceptualization, Methodology, Writing – review & editing. **Rodrigo Jensen:** Resources, Writing – review &&editing.

Data Availability

Data will be made available on request. The protocol of this study was registered in the International Prospective Register of Ongoing Systematic Reviews PROSPERO database (www.crd.york.ac.uk/prospero; registration number CRD42022303019).

Acknowledgments

The authors wish to thank Silvia Arranz Ferré and Concepcion Gonzalez Nieves, librarians at University of Valencia, for their contribution to the database search, and Margarete Penteado for the English language review.

Declarations of Interest

None.

Data profile - Data sets

Author: fenasta@alumni.uv.es.
Author not specifying any data sets in this Data Profile.

Reason

The data that has been used is confidential.

Comments

The raw/processed data required to reproduce the above findings cannot be shared at this time as the data also forms part of an ongoing study.

References

References marked with an asterisk indicate studies that were included in this research's systematic review.

Ainsworth, M. D. S., Bell, S. M., & Stayton, D. F. (1974). Infant-mother attachment and social development: Socialization as a product of reciprocal responsiveness to signals. In M. P. M. Richards (Ed.), *The integration of a child into a social world* (pp. 99–135). Cambridge University Press.

Ainsworth, M. D. S., Blehar, M. C., Waters, E., & Wall, S. (1978). *Patterns of attachment: A psychological study of the strange situation*. Lawrence Erlbaum.

Ainsworth, M. S. (1969). Object relations, dependency, and attachment: A theoretical review of the infant-mother relationship. *Child Development*, 969–1025.

Ainsworth, M. S. (1979). Infant–mother attachment. *American Psychologist, 34*(10), 932. https://doi.org/10.1037/0003-066X.34.10.932

Ainsworth, M. S., & Bowlby, J. (1991). An ethological approach to personality development. *American Psychologist, 46*(4), 333–341. https://doi.org/10.1037/0003-066X.46.4.333

Alvarenga, P., Kuchirko, Y., Cerezo, M. A., Mendonça Filho, E. J., Bakeman, R., & Tamis-LeMonda, C. S. (2021). An intervention focused on maternal sensitivity enhanced mothers' verbal responsiveness to infants. *Journal of Applied Developmental Psychology, 73*, Article 101313. https://doi.org/10.1016/j.appdev.2021.101313

(*) Anderson, G. C., Radjenovic, D., Chiu, S. H., Conlon, M., & Lane, A. E. (2004). Development of an observational instrument to measure mother-infant separation post birth. *Journal of Nursing Measurement, 12*(3), 215–234. https://doi.org/10.1891/jnum.12.3.215.

**Baird, S. M., Haas, L., McCormick, K., Carruth, C., & Turner, K. D. (1992). Approaching an objective system for observation and measurement: Infant-Parent social interaction code. *Topics in Early Childhood Special Education, 12*(4), 544–571. https://doi.org/10.1177/027112149201200410

**Bakeman, R., & Brown, J. V. (1977). Behavioral dialogues: An approach to the assessment of mother-infant interaction. *Child Development, 48*(1), 195. https://doi.org/10.2307/1128898

**Barnard, K. E., Hammond, M. A., Booth, C. L., Bee, H. L., Mitchell, S. K., & Spieker, S. J. (1989). Mesurement and meaning of parent–child interaction. *Psychological Development in Infancy*, 39–80. https://doi.org/10.1016/B978-0-12-041203-7.50007-4

Byrne, M. W., & Keefe, M. R. (2003). Comparison of two measures of parent-child interaction. *Nursing Research, 52*(1), 34–41.

Bowlby, J. (1969). Attachment and loss. Random House.

Bowlby, J. (1982). Attachment and loss: Retrospect and prospect. *American Journal of Orthopsychiatry, 52*(4), 664.

**Brazelton, T.B. (1973). Neonatal behavioral assessment scale. J. B. Lippnicott Co.

Brazelton, T. B., & Nugent, J. K. (1995). *Neonatal behavioral assessment scale*. Cambridge University Press.

**Censullo, M., Bowler, R., Lester, B., & Brazelton, T. B. (1987). An instrument for the measurement of infant-adult synchrony. *Nursing Research, 36*(4), 244–248. https://doi.org/10.1097/00006199-198707000-00015

Cerezo, M. A. (1995). The psychological impact of abuse: Early childhood and school age. *Infancia York Aprendizaje, 18*(71), 135–157. https://doi.org/10.1174/02103709560575541

Cerezo, M. A., Abdelmaseh, M., Trenado, R. M., Pons-Salvador, G., & Bohr, Y. (2021). The temporal dimension in the understanding of maternal sensitivity in caregiver-infant interactions: The 'Early Mother-Child Interaction Coding System'. *Infant Behavior and Development, 63*, Article 101563. https://doi.org/10.1016/j.infbeh.2021.101563

**Costas Moragas, C., Fornieles Deu, A., Botet Mussons, F., Boatella Costa, E., & de Cáceres Zurita, M. L. (2007). Evaluación psicométrica de la Escala de Brazelton en una muestra de recién nacidos españoles. *Psicothema, 19*(1), 140–149. ⟨http://www.psicothema.com/pdf/3340.pdf⟩.

**Dumas, L., Lepage, M., Bystrova, K., Matthiesen, A. S., Welles-Nyström, B., & Widström, A. M. (2013). Influence of skin-to-skin contact and rooming-in on early mother–infant interaction. *Clinical Nursing Research, 22*(3), 310–336. https://doi.org/10.1177/1054773812468316

Eickmann, S. H., Emond, A. M., & Lima, M. (2016). Evaluation of child development: beyond the neuromotor aspect. *Jornal Délelőtt Pediatria, 92*, 71–83.

**Feldman, R. (1998). Coding Interactive Behavior (CIB). Center for Attachment Research. https://center-for-attachment.com/portfolio-item/coding-interactive-behavior-cib/.

Feldman, R., Eidelman, A. I., Sirota, L., & Weller, A. (2002). Comparison of skin-to-skin (kangaroo) and traditional care: Parenting outcomes and preterm infant development. *Pediatrics, 110*(1), 16–26.

Galván-Bovaira, M. J., Gràcia, M., & del Río, M. J. (2002). A proposal for the evaluation of mother-child communicative interactions: The mother-infant communication screening (MICS). *Revista de Logopedia. Foniatría York Audilogía, 22*(1), 15–23. https://doi.org/10.1016/s0214-4603(02)76217-7

Guida, E., Scano, A., Storm, F., Biffi, E., Reni, G., & Montirosso, R. (2021). Mother-Infant Interaction Kinect Analysis (MIIKA): An automatic kinematic-based methodology for the investigation of interpersonal distance during early exchanges. *Infant Behavior and Development, 63*, Article 101567. https://doi.org/10.1016/j.infbeh.2021.101567

**Heimann, M. (1989). Neonatal imitation, gaze aversion, and mother-infant interaction. *Infant Behavior and Development, 12*(4), 495–505. https://doi.org/10.1016/0163-6383(89)90029-5

Hernández, C., Muñoz-Ledo, R., Figueroa, O., Rivera, G., Méndez, R. & Sánchez, P. (2013). Concordance between two instruments for assessment of the early mother-child interactions. Ciencias Clínicas, 14(2), 45–50. https://www.elsevier.es/es-revista-ciencias-clinicas-399-articulo-concordancia-entre-dos-instrumentos-evaluacion-X1665138314356672.

Høifødt, R. S., Nordahl, D., Landsem, I. P., Csifcsák, G., Bohne, A., Pfuhl, G., Rognmo, K., Braarud, H. C., Goksøyr, A., Moe, V., Slinning, K., & Wang, C. E. A. (2020). Newborn behavioral observation, maternal stress, depressive symptoms and the mother-infant relationship: Results from the northern babies longitudinal study (NorBaby. *BMC Psychiatry, 20*(1), 300. https://doi.org/10.1186/s12888-020-02669-y

Horowitz, J. A., Logsdon, M. C., & Anderson, J. K. (2005). Measurement of maternal-infant interaction. *Journal of the American Psychiatric Nurses Association, 11*(3), 164–172. https://doi.org/10.1177/1078390305278788

Hou, Y., Tian, J., Zhang, J., Yun, R., Zhang, Z., Chen, K. H., Zhang, C., & Wang, B. (2017). Quality of meta-analysis in nursing fields: An exploration based on the JBI guidelines. Article e0177648 *PLoS One, 12*(5). https://doi.org/10.1371/journal.pone.0177648.

Ionio, C., Ciuffo, G., & Landoni, M. (2021). Parent–infant skin-to-skin contact and stress regulation: A systematic review of the literature. *International Journal of Environmental Research and Public Health, 18*(9), 4695.

**Kumar, R., & Hipwell, A. E. (1996). Development of a clinical rating scale to assess mother-infant interaction in a psychiatric mother and baby unit. *British Journal of Psychiatry, 169*(1), 18–26. https://doi.org/10.1192/bjp.169.1.18

Lourenço, V., Coutinho, J., & Pereira, A. F. (2021). Advances in microanalysis: Magnifying the social microscope on mother-infant interactions. *Infant Behavior and Development, 64*, Article 101571. https://doi.org/10.1016/j.infbeh.2021.101571

Lotzin, A., Lu, X., Kriston, L., Schiborr, J., Musal, T., Romer, G., & Ramsauer, B. (2015). Observational tools for measuring parent–infant interaction: A systematic review. *Clinical Child and Family Psychology Review, 18*(2), 99–132. https://doi.org/10.1007/s10567-015-0180-z

Mesman, J., & Emmen, R. A. (2013). Mary Ainsworth's legacy: A systematic review of observational instruments measuring parental sensitivity. *Attachment & Human Development, 15*(5–6), 485–506.

Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., Bouter, L. M., & De Vet, H. C. (2010). The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: An international Delphi study. *Quality of Life Research, 19*(4), 539–549. https://doi.org/10.1007/s11136-010-9606-8

Morton, S. U., & Brodsky, D. (2016). Fetal physiology and the transition to extrauterine life. *Clinics in Perinatology, 43*(3), 395–407. https://doi.org/10.1016/j.clp.2016.04.001

Munson, L. J., & Odom, S. L. (1996). Review of rating scales that measure parent-infant interaction. *Topics in Early Childhood Special Education, 16*(1), 1–25. https://doi.org/10.1177/027112149601600104

National Scientific Council on the Developing Child (2020). Connecting the brain to the rest of the body: early childhood development and lifelong health are deeply intertwined. https://developingchild.harvard.edu/resources/connecting-the-brain-to-the-rest-of-the-body-early-childhood-development-and-lifelong-health-are-deeply-intertwined/.

Naughton, A., Perkins, L., McMinn, B., & Kemp, A. (2019). Using an observation tool (Parent–Infant Interaction Observation Scale) to assess parent–infant interaction in the first 2 weeks of life: A feasibility study. *Child: Care, Health and Development, 45*(2), 271–285.

Norholt, H. (2020). Revisiting the roots of attachment: A review of the biological and psychological effects of maternal skin-to-skin contact and carrying of full-term infants. *Infant Behavior and Development, 60*, Article 101441. https://doi.org/10.1016/j.infbeh.2020.101441.

Ouzzani, M., Hammady, H., Fedorowicz, Z., & Elmagarmid, A. (2016). Rayyan - a web and mobile app for systematic reviews. *Systematic reviews, 5*, 1–10.

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, McDonald, S., & Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ, 372*(71), 1–9. https://doi.org/10.1136/bmj.n71

Pasquali, L. (2001). Técnicas de Exame Psicológico–TEP: manual. Casa do Psicólogo.

Perone, S., Gartstein, M. A., & Anderson, A. J. (2020). Dynamics of frontal alpha asymmetry in mother-infant dyads: Insights from the Still Face Paradigm. *Infant Behavior and Development, 61*, Article 101500. https://doi.org/10.1016/j.infbeh.2020.101500

Polit, D.F. & Beck, C.T. (2019). Fundamentos de pesquisa em enfermagem: Avaliação de evidências para a prática da enfermagem. Artmed Editora.

**Provenzi, L., Olson, K., Giusti, L., Montirosso, R., Desantis, A., & Tronick, E. (2018). NICU Network Neurobehavioral Scale: 1-month normative data and variation from birth to 1 month. *Pediatric Research, 83*(6), 1104–1109. https://doi.org/10.1038/pr.2018.25

Raack, C.B. (1989). Mother/Infant Communication Screening (MICS). Roselle, IL: Community Therapy Services.

Rattaz, V., Puglisi, N., Tissot, H., & Favez, N. (2022). Associations between parent–infant interactions, cortisol and vagal regulation in infants, and socioemotional outcomes: A systematic review. *Infant Behavior and Development, 67*, Article 101687. https://doi.org/10.1016/j.infbeh.2022.101687

Reeb-Sutherland, B., Williams, L. R., Gartstein, M. A., & Fox, N. A. (2022). Methodological advances in the characterization and understanding of caregiver-infant interactions. *Infant Behavior and Development, 66*, Article 101668. https://doi.org/10.1016/j.infbeh.2021.101668

**Rime, J., Tissot, H., Favez, N., Watson, M., & Stadlmayr, W. (2018). The diaper change play: Validation of a new observational assessment tool for early triadic family interactions in the first month postpartum. *Frontiers in Psychology, 9*, 497. https://doi.org/10.3389/fpsyg.2018.00497

Rosenberg, S. A., Robinson, C. C., & Beckman, P. J. (1986). Measures of parent–infant interaction: An overview. *Topics in Early Childhood Special Education, 6*(2), 32–43. https://doi.org/10.1177/027112148600600204

Rosenkoetter, U., & Tate, R. (2018). Assessing features of psychometric assessment instruments: a comparison of the COSMIN checklist with other critical appraisal tools. *Brain Impairment, 19*(1), 103–118. https://doi.org/10.1017/BrImp.2017.29

Sadeh, A., & Acebo, C. (2002). The role of actigraphy in sleep medicine. *Sleeping Medicine Reviews, 6*(2), 113–124. https://doi.org/10.1053/smrv.2001.0182

*Santos, M. S. (2008). Construcción de una escala de observación del vínculo madre-bebé internado en UCIN: Resultados preliminares de la aplicación del protocolo de observación. *Anuario Délelőtt Investigaciones, 15*, 207–217.

**Santos, M.S. (2010). Psychoneonatology: Results of a new observation scale of the bond mother-baby in incubator. In A. Puertas, F. Montoya, J. Romero, J. Hurtado, S. Manzanares (Eds.), Advances in perinatal medicine (pp. 515–519). Monduzzi editore. http://hdl.handle.net/11336/110361.

Santos, W. M. D., Secoli, S. R., & Püschel, V. A. D. A. (2018). The Joanna Briggs Institute approach for systematic reviews. *Revista Latino-Americana Délelőtt Enfermagem, 26*, 3074. https://doi.org/10.1590/1518-8345.2885.3074

**Silberstein, D., Feldman, R., Gardner, J. M., Karmel, B. Z., Kuint, J., & Geva, R. (2009). The mother-infant feeding relationship across the first year and the development of feeding difficulties in low-risk premature infants. *Infancy, 14*(5), 501–525. https://doi.org/10.1080/15250000903144173

Svanberg, P. O., Barlow, J., & Tigbe, W. (2013). The Parent-Infant Interaction Observation Scale: Reliability and validity of a screening tool. *Journal of Reproductive and Infant Psychology, 31*(1), 5–14. https://doi.org/10.1080/02646838.2012.751586

Terwee, C. B., Jansma, E. P., Riphagen, I. I., & de Vet, H. C. (2009). Development of a methodological PubMed search filter for finding studies on measurement properties of measurement instruments. *Quality of Life Research, 18*(8), 1115–1123. https://doi.org/10.1007/s11136-009-9528-5

Terwee, C. B., Bot, S. D., de Boer, M. R., van der Windt, D. A., Knol, D. L., Dekker, J., Bouter, L. M. J., & de Vet, H. C. (2007). Quality criteria were proposed for measurement parameters of health status questionnaires. *Journal of Clinical Epidemiology, 60*(1), 34–42. https://doi.org/10.1016/j.jclinepi.2006.03.012

Trenado R.M. (2001). Potencial de abuso físico e interacción temprana madre-hijo con las conductas de apego en el primer año de vida [Doctoral dissertation, Universitat de València]. ProQuest Dissertations & Theses Global.

Tronick, E., Als, H., & Brazelton, T. B. (1980). Monadic phases: A structural descriptive analysis of infant-mother face-to-face interaction. *Merrill-Palmer Quarterly of Behavior and Development, 26*(1), 3–24. ⟨https://www.jstor.org/stable/23084074⟩.

White-Traut, R. C., Nelson, M. N., White-Traut, R. C., & Nelson, M. N. (1988). Maternally administered tactile, auditory, visual, and vestibular stimulation: Relationship to later interactions between mothers and premature infants. *Research in Nursing & Health, 11*(1), 31–39. https://doi.org/10.1002/nur.4770110106

Zirpoli, D. B., Mendes, R. B., Reis, T. S., Barreiro, M. D. S. C., & de Menezes, A. F. (2019). Benefits of the Kangaroo Method: An Integrative Literature Review. *Revista Délelőtt Pesquisa Cuidado é Fundamental Online, 11*(2), 547–554.