Infant Behavior and Development 78 (2025) 102017

Contents lists available at ScienceDirect

Infant Behavior and Development

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

Registered report stage ii

Toward a global understanding of neonatal behaviour: adaptation and validation of the Neonatal Behavioural Assessment Scale (NBAS) in the UK and rural Gambia

Laura Katus[a,b,*], Maria Rozhko[c], Christine Torrance[d], Tijan Fadera[e], Fabakary Njai[e], Marta Perapoch Amadó[f], Bosiljka Milosavljevic[g], Samantha McCann[e,h], Mustapha Minteh[e], Malang Jammeh[e], Jane Barlow[i], Clare E. Elwell[j], Sophie E. Moore[e,h], Sarah Lloyd-Fox[c]

[a] *Institute for Lifecourse Development, University of Greenwich, UK*
[b] *Centre for Family Research, University of Cambridge, UK*
[c] *Department of Psychology, University of Cambridge, UK*
[d] *Department of Health Sciences, Warwick Medical School, University of Warwick, UK*
[e] *Medical Research Council Unit The Gambia at the London School of Hygiene and Tropical Medicine, UK*
[f] *Department of Psychology, University of East London, UK*
[g] *Department of Biological and Experimental Psychology, Queen Mary University of London, UK*
[h] *Department of Women and Children's Health, Kings College London, UK*
[i] *Department of Social Policy and Intervention, University of Oxford Centre for Evidence-Based Medicine, Oxford, UK*
[j] *Department of Medical Physics and Biomedical Engineering, UK*

ARTICLE INFO

Keywords:
Neonatal hebaviour
Neonatal Behavioural Assessment Scale (NBAS)
Cultural adaptation
Perinatal risk factors
The Gambia
UK

ABSTRACT

There is a need of expanding research on neonatal behaviour to encompass diverse global populations. However, few measures appropriate for use from birth in diverse cultural contexts exist. We present data from rural Gambia and the UK using the Neonatal Behavioural Assessment Scale (NBAS). In Phase 1, the scale was piloted for use in The Gambia, highlighting great utility for use in this setting. Adaptations included 1) additional explanation of some items to caregivers and 2) omission of items where the home environment necessitated to do so. In Phase 2, the NBAS was administered in both the UK and The Gambia. Item level comparisons across the sites showed fewer state changes in Gambian infants and a greater degree of examiner facilitation throughout the session. Factor analysis within the larger Gambian cohort indicated best model fits where first-order factors mapping onto each NBAS subscale were accompanied by a second-order 'Self-Organising System'-factor, mirroring prior factor analytic studies on the scale. Further, the habituation subscale had to be omitted from analyses due to large amounts missing data, highlighting potential differences across subscales when implementing the NBAS across diverse settings. We found associations between our NBAS factor scores and NBAS supplementary items. Examining known risk factors for early neonatal development, we found associations of the Social Interactive and Self-Organising System factor with pregnancy anxiety and gestational age at birth and birthweight, but not maternal anaemia. Our findings inform future studies seeking to understand the interplay between cultural contexts, perinatal factors, and early neurobehavioural development.

* Correspondence to: School of Human Sciences, University of Greenwich, Dreadnaught Building, Park Row, London SE10 9LS, UK.
E-mail address: l.katus@greenwich.ac.uk (L. Katus).

https://doi.org/10.1016/j.infbeh.2024.102017
Received 6 June 2024; Received in revised form 24 November 2024; Accepted 30 November 2024
Available online 4 December 2024

1. Introduction

1.1. Mental well-being begins in infancy

An infant's ability to experience, express and regulate emotions, form secure relationships, explore, and learn from the environment, provides an important foundation for life-long development (Clinton et al., 2016; Zeanah, 2018). The nature of infants' early relationships and caregiver attachment have been associated with a host of outcomes, including neurocognitive (Puhlmann et al., 2021) and socio-emotional development (Groh et al., 2017), as well as long-term mental health (Galbally et al., 2020). Development and wellbeing during the neonatal period is shaped by a reciprocal interplay with caregivers and the wider social environment and has been shown to be adversely affected by poverty-related risk factors (Sameroff et al., 2004). While links between newborn behaviour, socio-economic and psychosocial risk factors (e.g., perinatal parental mental illness (Gressier et al., 2020) have long been recognised (Richter, 2003), these insights have not yet been consistently translated into changes in practice. Over 20 years ago, Sander (2000) highlighted the need to think about well-being during the neonatal period as a catalyst for long-term health and well-being, both on an individual and a societal level. Specifically, it was highlighted that there is a need to gather high-quality empirical data from across the world to better understand the way in which environmental risk and protective factors come together to impact infants perinatally during a sensitive developmental period. Contributing to this line of research, this study examines newborn mental health via the Neonatal Behavioural Assessment Scale (NBAS, Brazelton and Nugent, 2011) in two contrasting settings; rural Gambia, West Africa and in the UK. We lay out adaptations made to the NBAS for use in rural Gambia, before examining its psychometric properties in the Gambian context, where neontatal mental health has never previously been examined. Lastly, we assess how perinatal infant (birthweight and gestational age) and maternal (anxiety, pregnancy anaemia) factors may be mechanistically associated with infant NBAS scores.

1.2. Assessing neonatal behaviour and development: associations with environmental risk factors and long-term outcomes

In part, the current lack of global research within the field of neonatal development can be attributed to a lack of measuring tools. As reviewed by Szaniecki and Barnes (2016), assessments currently available include questionnaire-based measures (e.g., Child Behaviour Checklist [18 + months, Achenbach & Rescorla, 2000], Infant-Toddler Symptom Checklist [7 + months, DeGangi, et al., 1995], Infant-Toddler Social and Emotional Assessment, as well as its abbreviated version [12 + months, Carter et al., 2003, Briggs-Gowan, et al., 2004], Ages and Stages Questionnaires-Social-Emotional version [from birth, Squires et al., 2002]) as well as structured and unstructured observational tools. In contrast to most questionnaire-based measures which cannot be implemented from birth, behavioural observational tools exist that are well-validated and can be used even in the youngest age groups. Amongst these, two of the most used structured observational assessments are the NBAS, and its shorter counterpart the Neonatal Behavioural Observation (NBO, Brazleton & Nugent, 1997, Brazleton & Nugent, 2011). The NBAS is a comprehensive tool measuring behavioural and neurological responses of newborn infants. Typically conducted within the first few days after birth (although it can be administered up to 8 weeks of age), it involves observing the infant's interactions with their environment. The scale provides a standardised framework to evaluate various aspects of an infant's behaviour, including motor skills, reflexes, as well as their orientation and responsiveness to stimuli (e.g., orientation to faces, voices, or non-social objects).

NBAS responses have been longitudinally associated with developmental outcomes. Shoaff et al. (2021) showed that NBAS measures were associated with socio-communicative abilities in adolescence in a US sample; specifically, self-regulation, habituation to auditory stimuli, and autonomic nervous system regulation at 2 weeks were linked to social communication at 15 years of age. Canals et al. (2006) highlighted associations of orientation and habituation responses with externalising and internalising problems at 6 years of age in a Spanish sample. Within early childhood, the association between early NBAS responses and social-communication outcomes has also been demonstrated: in a study of 105 Japanese infants, Tokunga et al. (2019) found that NBAS scores at 2–7 days of age differed for those infants with and without clinical scores on a checklist of autism spectrum disorder at 18 months of age. The NBAS has also been found to be sensitive to rapid neurodevelopmental changes occurring perinatally as infants' self-regulatory abilities were associated with number of weeks' gestation at birth (Malak et al., 2021).

Responses on the NBAS/NBO have further shown to have robust associations with several relevant environmental risk exposures. There is a well-documented link between NBAS responses and maternal mental health: in a study of 598 mother-infant pairs, Gressier and colleagues (2020) found that prenatal depressive symptoms were independently associated with neonatal measures of habituation, orientation, motor system and autonomic stability, controlling for both postnatal depression and anxiety. Similarly, lower NBAS motor (Zhang et al., 2019) and social communication (Osborne et al., 2022) scores were identified in infants of perinatally depressed mothers (Zhang et al., 2019). Furthermore, links have been established between maternal anxiety and infant irritability on the NBO, with maternal mental health support leading to improved scores (Zhang et al., 2018). Based on these prior findings, as well as recent investigations by our own group highlighting the prevalence of maternal stress and anxiety in The Gambia (Milosavljevic et al., under review), we examined associations between maternal mental health and NBAS scores in the current study.

Poorer NBAS scores have also been reported in at-risk populations (e.g., exposure to environmental toxins, Stewart et al., 2000, Sagiv et al., 2008, Kooistra et al., 2006), as well as in the context of maternal iron deficiency (Hernández-Martínez et al., 2011), preterm birth (Greene, Fox, Lewis, 1983, Alvarez-Garcia et al., 2015) and small for gestational age infants (Figueras et al., 2009). In reference to this literature, as well as an increased prevalence of maternal anaemia (McCann et al., 2023) and low infant birthweight in The Gambia (Hennig et al., 2017), the current study investigated links between these risk factors and NBAS scores.

The NBAS has been implemented across a range of countries. Due to the nature of the outcomes measured, it is considered to have universal or transcultural application (Brazelton & Nugent, 2011) and to date has been widely used in diverse settings world-wide (Fig. 1).

Here, we are expanding on this work by providing information on the NBAS' first use in West Africa and only the third from sub-Saharan Africa. Of note, only a subset of previous studies reported on parental perceptions of and how these necessitated adaptations of the scale; however, studies that included such information highlighted some key differences. When using the NBAS across Kenya and the US, Super & Harkness, (2020a,2020b) reported that US mothers focussed more on aspects that might inform their infant's cognitive development, whereas mothers in Kenya placed greater emphasis on motor responses. As differing perceptions of the NBAS items may be associated with parental acceptability in The Gambia, we conducted a pilot study to test the scale's acceptability, assessing caregiver perception of the scale being carried out with their infant, before examining its psychometric properties in more depth. Since the NBAS has been widely used in UK settings (Hill et al., 2013; Gunning et al., 2013; Cecil et al., 2017; Froggatt et al., 2020; Osborne et al., 2022), no such pilot was conducted in our UK cohort.

1.3. NBAS structure and measurement properties

The comprehensive NBAS protocol consists of items assessing both infant behaviour and reflexes, administered by a qualified examiner. Based on observed behavioural and physiological responses (e.g., skin colour changes, startles, tremors, respiration) and state changes (e.g., transitions from calm alert to crying/fussing states), 27 behavioural, 18 reflex and 7 supplementary items (i.e., summary items designed to provide a more qualitative assessment) are scored. While responses sometimes are examined on an item-by-item level (Brazleton et al., 1976, Tronick et al., 1976), this approach is prone to chance results due to the large number of statistical tests performed on intercorrelated measures. Therefore, data reduction to determine either an a-priori or an a-posteriori defined subset of response clusters is commonly applied.

The cluster scores are calculated as the mean of item scores comprising a particular cluster, with higher cluster scores indicating better performance. Items with a mid-range optimal score are re-coded as linear items before contributing to the cluster score. The most widely used method is the original factors drawn up by Lester and colleagues (1982, 1984), which group items into seven clusters: (i) Habituation, measuring response decrements to repetitive visual, auditory and tactile stimulation during sleep; (ii) Orientation/Social Interaction, examining infants' gaze and movement towards visual and auditory stimuli, and the quality of overall alertness; (iii) Motor System, measuring motor performance and the quality of movement and muscle tone; (iv) State Organisation, which evaluates the quality and stability of the infant's states, providing insight into their overall physiological organisation and adaptability; (v) State Regulation, measuring how the infant responds to environmental stimuli (provided by the examiner) and the ability to manage their

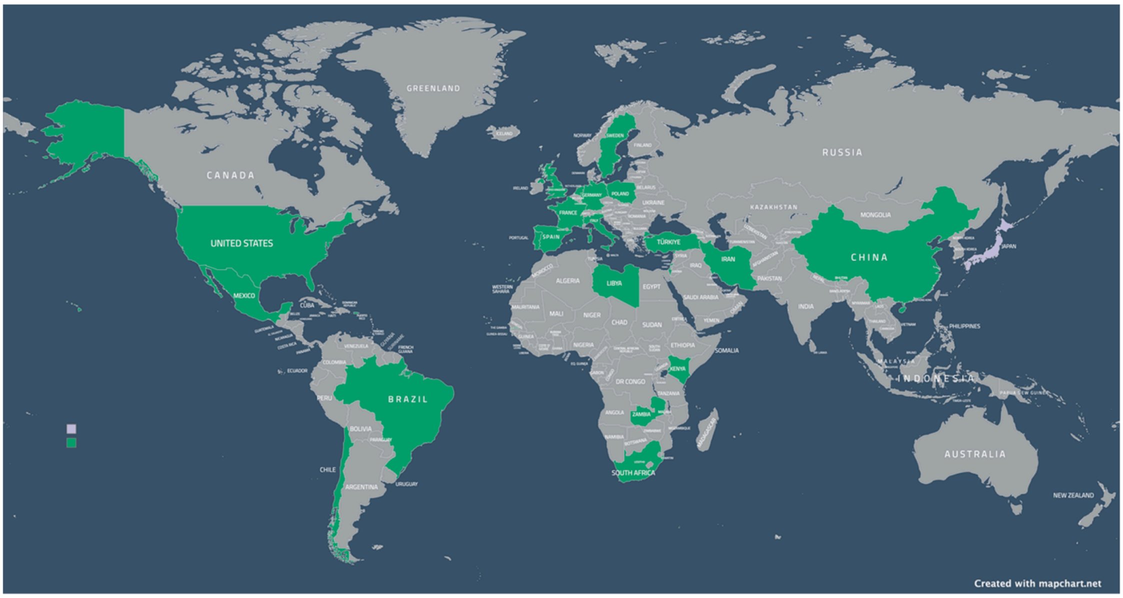

Fig. 1. Countries in which the NBAS has to date been implemented. These include Chile (Ayala et al., 2021), China (Wang et al., 2022, Zhang et al., 2020; 2019; 2018; 2017; Chen et al., 2017), France (Belot et al., 2021), Germany (Grossmann et al., 1985), Iran (Aliabadi & Askary, 2013), Israel (Tirosh et al., 1992; Lowinger, 1999; Feldman & Eidelman, 2006), Italy (Purpura et al., 2023), Japan (Eishima, 1992, Ohgi et al., 2003, Loo et al., 2005, Tokunaga et al., 2019), Kenya (Super & Harkness, 2020a, 2020b), Mexico (Soler-Limón et al., 2019), Poland (Tabaczyńska et al., 2021; Malak et al., 2022), Portugal (Barbosa et al., 2018), Puerto Rico (Coll et al., 1982), South Africa (Rencken et al., 2022), Spain (Canals et al., 2006; 2011, Costas Maragas et al., 2007; Costa et al., 2010; Alvarez-Garcia et al., 2015; Ayesa-Arriola et.al., 2023), Sweden (Lundqvist-Persson et al., 2012), the Netherlands (Wolf et. al., 2002), the UK (Hill et al., 2013; Gunning et al., 2013; Cecil et al., 2017; Froggatt et al., 2020; Osborne et al., 2022), Turkey (Basdas et al., 2018) and Zambia (Stengel, 1981)

arousal; (vi) Autonomic Stability, noting any startles and marked changes in skin tone during the session; and (vii) Reflexes, which notes the number of abnormal elicited responses (Lester, Als & Brazelton, 1982). When originally designed, these seven clusters were found to provide the best model fit for the data available, and various studies have employed this model to analyse infant behaviour since (Feldman, 2006, 2015; Sagiv et al., 2008; Canals et al., 2011; Tabaczyńska et al., 2021; Rencken et al., 2022; Osborne et al., 2022).

Since the Lester factors were conceptually rather than empirically derived, a separate branch of studies has set out to examine their psychometric properties across different study populations. In N = 166 US preterm infants (Azuma et al., 1991), confirmatory factor analysis (CFA) showed an inadequate fit of the Lester model due to high intercorrelations of factors, contrary to the original findings of relative cluster independence. Azuma et al. (1991) suggested that these differences and the lack of fit of the Lester model might be due to the specific characteristics of the cohort. Nevertheless, the Lester factors continued to be the primary data reduction method in clinical studies (Wolf et al., 2002; Ohgi et al., 2003; Medoff-Cooper & Ratcliffe, 2005; Feldman & Eidelman, 2006; Tabaczyńska et al., 2021; Malak et al., 2022). More recent investigations of the psychometric properties of the Lester factors using NBAS data collected with term-born infants have also found an overall inadequate fit of the model. To test a theory that such inconsistency might arise from some factors having large proportions of missing items and therefore reducing the amount of available data for some factors, (e.g., Azuma et al. (1991) excluded Habituation and Reflex items due to missing data), McCollam et al. (1997) performed CFA on 160 healthy term-born infants that had data for all NBAS items. Firstly, they tested the fit of the Lester model when the factors were allowed or not allowed to correlate, showing a significantly better overall fit of the correlated model, although both models showed an inadequate overall fit. The authors proposed extending the Lester six-factor model by allowing the items within each factor to load onto a seventh 'General Adaptiveness' factor, which yielded a significantly better fit than the original six-factor model. More recently, Barbosa and colleagues (2018) performed CFA on NBAS data collected with a sample of 196 term-born Portuguese infants, proposing several modifications to the scoring and the utility of the original items. In addition to omitting Reflexes and Smiles from analysis, they proposed to exclude items of Cuddliness, Motor Maturity, Defensive Movements, Response Decrement to Tactile Stimulation of the Foot and Rapidity of Build-up items from the analysis as these were found to substantially reduce model fit. Models were tested with two primary aims: (1) to compare Lester's behavioural cluster system, tested with original NBAS scoring, with a proposed model based on Barbosa et al.'s modifications to administration and scoring; (2) to compare a model where first-order factors were allowed to correlate with a model where factor correlations were explained by a second-order 'Self-Organising System' latent factor. The results indicated that the best model fit was achieved when the six Lester factors were obtained from revised rather than original item scoring and extended by the inclusion of a second order factor, in line with the findings of McCollam et al. (1997). In summary, the psychometric properties of the Lester model have raised questions about its applicability. Studies with premature infants and term-born infants have shown inconsistent fits with the model. In part, these differences may be due to population characteristics, such as infants' gestational age or age at testing. However, other factors such as missing data, re-coding of mid-range optimal score items, exclusion of items and item correlations have all posed challenges in confirming the model's adequacy. Efforts to improve the model's fit have led to modifications and extensions. Researchers have proposed additional factors, such as a 'General Adaptiveness' factor and a 'Self-Organizing System' latent factor, to account for cross-subscale item correlations. These modified models have shown a better overall fit than the original Lester model, emphasizing the importance of refining and adapting assessment tools to better capture infant behaviour.

1.4. The current studies

The current study uses data from two infant cohorts in the UK and The Gambia, and consists of two phases; a pilot and adaptation phase (Phase 1, Gambia only) and a cohort analysis phase (Phase 2, UK and Gambia). As the NBAS had never been used in West Africa prior to this, Phase 1 examined caregivers' responses and feedback to observing the scale, to be able to adapt items of the NBAS as appropriate for this setting. Phase 2 took place as part of the Brain Imaging for Global Health (BRIGHT) project (see Lloyd-Fox et al., 2024; Katus et al., 2019), which longitudinally followed infant cohorts from birth in the UK (N = 61) and The Gambia (N = 204). Prior work in this region of The Gambia, as well as by our group, has highlighted the prevalence of several contextual factors, such as maternal anaemia (McCann et al., 2023), low birth-weight (Hennig et al., 2017) and maternal anxiety (Milosavljevic et al., under review); risk factors previously associated with NBAS scores from other settings. Furthermore, developmental trajectories of early neural specialisation and cognitive development have been shown to differ markedly between the UK and Gambian cohorts (e.g., Lloyd-Fox et al., 2019; Katus et al., 2020). Here, we aim to expand on those findings, by examining associations of these known risk factors with responses occurring shortly after birth during the neonatal period, to test whether the effect of these exposures already takes effect during the first days and weeks of postnatal life.

Due to the limited sample size in the UK cohort, we only present item-level scores from this cohort alongside the Gambian cohort to highlight any potential differences before assessing the psychometric properties of the scale in more depth in the larger Gambian sample. We then proceed to conduct CFA based on the Gambian cohort. For Phase 2, we aimed to i) assess the model fit for the original Lester factors in the Gambian cohort, ii) assess an alternative second-order model fit suggested by Barbosa et al. (2018) to assess which of the two model provides a better fit for the data, and iii) assess associations of resulting factor scores with relevant perinatal infant (gestational age, birthweight) and maternal (gestational anaemia, maternal mental health) predictors.

2. Phase 1: Pilot Study

2.1. Method

2.1.1. Study Setting

Data was collected in West Kiang, a rural area of The Gambia comprising 36 villages within the Lower River Region (Hennig et al., 2015). Compared to the coastal, urban capital of Banjul, West Kiang is characterised by lower socio-economic status (SES), with most people supporting themselves through subsistence agriculture (Hennig et al., 2015). Villages are subdivided into compounds, within which reside family units of 1–170 people, with an average of 16 members (Hennig et al., 2015). Caregiving responsibilities are often divided between members of these larger family structures, and especially in later infancy encompass not only parents, but also aunts, grandmothers, and co-wives (Katus et al., 2024). In accordance with Mandinka customs, the mother and infant stay inside the home for a period of confinement, until a naming ceremony is held around the 8th day after birth.

2.1.2. Participants

Participants were identified via the Demographic Surveillance System and were eligible if they were above the age of 18 years and mothers, fathers or grandparents of an infant < 2 months of age. Participants provided informed consent during home visits. Only members of the majority ethnic group of Mandinka were eligible to avoid confounds of translating feedback from multiple languages.

2.1.3. Measures

2.1.3.1. NBAS. The NBAS (fourth edition, Brazelton & Nugent, 2011) is a neurobehavioural scale assessing newborn behaviour and individual differences in self-regulation during the period of adaptation to postnatal life. The NBAS assesses the newborn's behavioural repertoire within the dynamic context of the infant-caregiver relationship through 27 behavioural items, 7 supplementary items (both scored on a 9-point scale) and 18 reflex items (scored on a 4-point scale). The scale is suitable for use from hours after birth up to eight weeks of age (in context of full-term infants) or eight weeks' corrected age (in context of preterm infants) and takes approximately 30 minutes to complete.

2.1.3.2. NBAS feedback form. The NBAS feedback form was developed for the purposes of this pilot study with the goal of gathering information directly from parents of newborns, and their own parents in the previous generation, on: (i) parents' and grandparents' overall positive or negative impressions of the NBAS; and (ii) whether parents felt that each item they saw was acceptable, and, in their view, entailed only actions that were appropriate to perform on infants in their culture, and that parents would likely feel comfortable

Table 1
NBAS pilot study feedback form.

General feedback
1. Was there anything that upset you or made you uncomfortable?
2. Do you think there was anything strange about the session, or anything you did not agree with?
3. Do you think other mothers and fathers in Kiang West would find acceptable such a session with their babies?
Item-specific feedback
For each item I describe (….) please choose one of the following three categories: was it (good/interesting), (acceptable/fine), or (not acceptable/made you uncomfortable/ you did not agree).
Habituation Package
4. When your baby was sleeping, and we shook the rattle and rang the bell?
5. When your baby was sleeping, and we shined the light across his/her eyes?
6. When your baby was sleeping, and we undressed him/her and touched his/her foot with the probe?
Motor-Oral Package
7. When we looked at your baby's muscle tone in the arms and legs?
8. When I stroked your baby's cheek and placed my gloved finger in his/her mouth to look at the sucking response?
9. When I looked at your baby's hand grasp, and touched his/her feet with my finger, and pushed back on the foot toward the leg?
10. When I tapped your baby's forehead?
Truncal Package
11. When I undressed your baby?
12. When I pulled your baby up to sit from a lying position?
13. When I placed your baby's feet against the table one by one?
14. When we looked at your baby 'standing' and 'walking'?
15. When I ran my finger down his/her spine and we watched the hips swing?
16. When I placed your baby on his/her tummy?
17. When I held your baby facing me and spun to the left and right?
18. When I held your baby close to me? Vestibular Package
19. When I placed the cloth over your baby's eyes and we watched his/her response?
20. When I placed your baby's head to the left and to the right?
21. When I let your baby experience a short drop into my arms?
Social Interactive Package
22. When we watched your baby's response to the ball?
23. When we shook the rattle on the side of your baby's head, and when we watched his/her response to the rattle being held and shaken in front of him/her?

with if performed as part of a research study. The questions asked can be found in Table 1. Items assessing infant responses to the examiner's face and voice were not included in this questionnaire, as these are typical of infant engagement and handling and therefore were not expected to elicit specific responses from caregivers.

The first three questions broadly explored participants' observations and opinions of the session. Subsequent questions addressed the acceptability of individual NBAS items. A shortened version using the first 10 questions only was administered to grandparents to assess overall impressions. Feedback on individual NBAS items was only collected from the infants' parents. For the second part of the feedback form, mothers and fathers were instructed to evaluate each NBAS item or set of items as positive, neutral, or negative.

2.1.4. Procedure

NBAS sessions took place in participants' homes. Participants were reminded that all adults who would be providing feedback should observe the full assessment session, and that feedback would be audio recorded. All sessions were performed by a trained and certified NBAS administrator. In the case of multiple observers, all participants jointly observed the NBAS session but were interviewed individually to gather their subsequent feedback. Any running commentary during the session was recorded on the feedback form. Forms were read out in English, and simultaneously translated by one of two native Mandinka speakers, who translated back caregivers' feedback. The process of simultaneous translation constituted a less-standardized approach that, while providing important insight in this pilot study, was replaced by formal translation for Phase 2 of the project (see below). All caregivers (parents and grandparents) responded to the three general feedback questions.

2.1.5. Data analysis

Responses were recorded as categorical data (positive, neutral, negative) for each question. Participants were also asked to elaborate on their responses. These free-form comments were used to further contextualise the categorical feedback.

2.2. Results

2.2.1. Sample characteristics

A total of N = 14 infants and N = 27 of their family members were selected; specifically: fourteen mothers between the ages of 18 and 39 years (median=28.5), eight fathers between the ages of 30 and 80 years (median=44), and five grandparents (two female) between the ages of 52 and 90 years (median=70) took part in the pilot study. All respondents lived in one of three villages in West Kiang (specifically in the villages of Keneba, Tankular and Kuli Kunda) and all but one (born in Senegal) were born in The Gambia, were Muslim, and fluent in Mandinka. Four respondents reported formal education in an English School, with the remainder reporting education at an Arabic School (emphasizing religious studies). Fifteen respondents reported agricultural activities (farming, gardening and cattle herding) as their occupation. Other reported occupations included driving, carpentry, electrician work, business work, school attendance (for two of the younger mothers), childcare, and housework.

2.2.2. NBAS feedback on whole session

The behaviours and items seen in the NBAS were new for most observers and were considered acceptable by all but two respondents. These participants queried the items of habituating to light and undressing the infant; however, upon further probing indicated that they found these items novel or unfamiliar, but not negative. No respondents said that any items were disagreeable. The remaining participants stated that they liked what they saw in the session and that the elicitation of items involved actions that were 'normal' for infants or infant handling. One mother's response provided evidence that explaining the rationale of the session and what would happen, before beginning, is a crucial way of ensuring the parents feel happy with the session. Sixteen respondents believed other mothers and fathers would find the NBAS acceptable. The remaining eleven respondents, rather than thinking other parents would find it unacceptable, said that they could not speak for other parents, some of whom might accept such a session and others of whom might not. This served as a reminder for examiners not to assume a given parent finds all NBAS items acceptable, and trainees were instructed to explain the content of the session to parents beforehand; 'talk through' certain items; watch for any signs of discomfort in the parent; and invite comments and questions after the session. There were no apparent systematic differences in responses between mothers, fathers and grandparents.

2.2.3. NBAS item specific feedback

2.2.3.1. Habituation. Mothers and fathers (N = 22) provided item-specific feedback on the NBAS. For each item, they indicated whether they rated it as positive, neutral or negative. Habituation items were rated as positive or neutral by most parents. One parent rated the rattle and bell as negative; however, no further rationale was provided. Given that four respondents listed the bell or rattle during sleep as their favourite item during the first part of the feedback session, and that the other 14 parents who saw this item gave positive (N = 12) or neutral (N = 2) feedback, no adjustment was recommended. Three parents rated the light as negative, indicating that light being shone in the face was 'not good', and one father (age 33 years) said that while he personally understood that the item was being done in the name of research and he therefore accepted it, "elderly ones don't want the light to be put on the face," seeming to point to a non-directly referenced cultural belief. As this was not corroborated by elders or other participants, and some had listed this item as their favourite, it was retained and explained prior to the habituation items starting. No parent commented negatively on the foot probe; however, one stated that "…when you used the probe on the foot, he moved, so that shows that the baby is well and very strong". It also needs to be noted that NBAS trainees are typically instructed to demonstrate the probe on their own and the parent's hand, to show that it is not painful.

2.2.3.2. Motor/reflex items. Most parents rated the reflex and motor items as positive or neutral. Most items (specifically passive tone in arms and legs; rooting and sucking; palmar, plantar, Babinski and ankle clonus; glabella; undressing; pull-to-sit; placing; standing and walking; incurvation; crawl; tonic deviation ("spin"); cuddliness; defensive; tonic neck reflex; and Moro) received only positive or neutral feedback. However, due to the frequent lack of suitable raised surfaces in participants' homes, assessors were instructed to only administer the pull-to-sit item where such a surface was available to minimise potential infant discomfort. Four items (undressing, placing, crawling, defensive) received isolated negative responses. In the case of crawling, one parent commented that the infant's umbilical cord was not fully healed: "There is no problem except for that of [baby], his umbilical cord is big, so for him that is the only problem when it comes to placing him on his tummy (Father, aged 55 years)". Therefore, trainees were advised to be aware of the umbilical stump when deciding whether to administer this item. For placing (stroking top of foot on bottom side of a surface to elicit extension of the leg), one parent commented that "[She] feels some pain, that's why she moved. (Mother, aged 18 years)". This item was however only administered to a small subset of infants (N = 4), due to the lack of even, raised surfaces with a soft ledge in many compounds, which is needed to elicit the reflex. The training therefore emphasised checking any ledges every time to make sure there was no danger of splintering and causing the infant pain, and if in doubt to use the NBAS manual as a make-shift surface to elicit the reflex. Undressing and the defensive items were found to be most controversial. One father was explicit in his evaluation of the undressing as negative, explaining: "…that one is not good because a newly born baby who is at this stage, not all the air is good for him so that is why we wrap babies in 2 or 3 cloths to prevent the air from reaching them. That's why to remove all the clothes is not good. It's just because it's part of your work that I allowed you to do that." (Father, aged 33 years). Based on this feedback, examiners were instructed to ask if parents were happy to have the infant undressed, and to skip the item or administer the reflex items in a vest if worn. The defensive item also elicited mixed feedback, with many parents asking for clarification prior to approving it (e.g., Mother, age 38 years: "That one, I don't understand, why did you do that? [examiner explains]. It's good."; Mother, age 20 years "Why did you do that? [examiner: To see his reaction – when the cloth is covering his eyes, what will he do? All babies are different, so it's to observe, for your child, if the cloth is placed there, what does he do?] There is no problem with that also."; Father, age 55 years "Putting the cloth on his face, that is to make the child to be curious, so when you put the cloth on his eyes, do you see any action? [examiner: What I observed is that when I put the cloth on the face, he shook his head.] Ok, there is no problem with that."). These responses led to examiners being instructed to provide an explanation when administering this item.

2.2.3.3. Orientation items and consoling manoeuvres. No negative feedback was given to the orientation or consoling items. However, some parents took away from the session that the items assessed the infants' hearing abilities. Examiners were thus instructed to make explicit that this test could not provide diagnostic information on the infant's hearing.

2.3. Study 1: Conclusion

Most items used in the NBAS were reviewed as positive or neutral by Gambian parents. Where negative comments were present, the protocol was adjusted to either i) provide additional explanation on the rationale for an item (e.g., for habituation, defensive) to check whether parents approved of the item's administration (e.g., undressing); or ii) allow examiners to skip an item if faced with an unsuitable environment (e.g., no suitable surface for pull to sit, placing reflex). Overall, the interviews emphasised the need to communicate with caregivers prior to and during the assessment, and in addition to observe the infant and be aware of cues from parents throughout. Another recurrent theme was that parents allowed certain actions to be performed specifically because they were part of a research study. This emphasised the ethical responsibility of making sure all items administered were relevant and provided a benefit to the participant or the wider community.

3. Phase 2: Implementation of the NBAS in the BRIGHT project

3.1. Method

3.1.1. Participants

For the Gambian site of the BRIGHT project, families were identified via the Demographic Surveillance System (DSS) with interested families providing informed consent during an antenatal clinic visit to the MRC Keneba field station (www.lshtm.ac.uk/research/units/mrc-gambia). Again, only Mandinka speakers were recruited. Participants were all residents of either the village of Keneba or surrounding villages in West Kiang. For the UK site, families were recruited during antenatal clinic visits to the Rosie Hospital, Cambridge University Hospitals NHS Foundation Trust. All families lived either in Cambridge or within a 20-mile radius. Participant numbers and reasons for attrition are provided in the below in the results section, for a more detailed breakdown please also see Lloyd-Fox, McCann, Milosavljevic, Katus et al., (2024).

3.1.2. Measures

3.1.2.1. UK. NBAS. The NBAS (fourth edition, Brazelton & Nugent, 2011) was administered when infants were between the ages of 7–14 days. Three trained and certified administrators conducted the study visits in participants' homes. After each session was run, assessors discussed and scored the session, with regular supervision being provided by a trained senior staff member.

3.1.2.2. Gambia. NBAS. The NBAS (fourth edition, Brazelton & Nugent, 2011) was administered as determined by the pilot study. For this phase of the study, two Gambian, Mandinka-speaking members of staff were trained and certified in the administration of the NBAS prior to data collection commencing. Initially, a trained assessor supported newly trained staff on site. In absence of a local NBAS trainer, staff video-recorded sessions as they completed their 20 practice assessments. These videos as well as their corresponding scoring sheets were then reviewed by a UK-based trainer, and regular feedback was provided. Assessors were then certified after submitting a final recorded session and individual scoring sheets.

Maternal and Infant perinatal factors. *Maternal Pregnancy Anxiety.* Pregnancy Anxiety was assessed during the third trimester via the Pregnancy Related Anxiety Scale (PRAS; Rini et al., 1999), which consists of 10 items asking participants to rate how frequently they had concerns about their pregnancy over the last few months, on a scale of 1–4 ("Not at all"-"A lot of the time"). Sum scores range from 10 to 40. While other questionnaires assessing maternal mental health were administered in the BRIGHT project, recent psychometric evaluation has highlighted the PRAS as the most robust in assessing maternal mental health, specifically pregnancy related anxiety, in this setting (Milosavljevic et al., under review).

Maternal Anaemia. Maternal anaemia was assessed via haemoglobin (Hb) concentrations (g/dL) measured in the third trimester of pregnancy from venous blood samples using a Medonic M-series 3- part Haematology Analyser (Boule Diagnostics AB, Sweden). Mothers were classified as anaemic using a criterion of 10.0 g/dL as an indicator of anaemia. This is the cut-off used by the Gambian healthcare system to administer treatment.

Infant gestational age and birthweight. Infant gestational age was estimated by ultrasound assessment at the antenatal booking scan. Infant birthweight was recorded during a neonatal examination within 72 hours of birth.

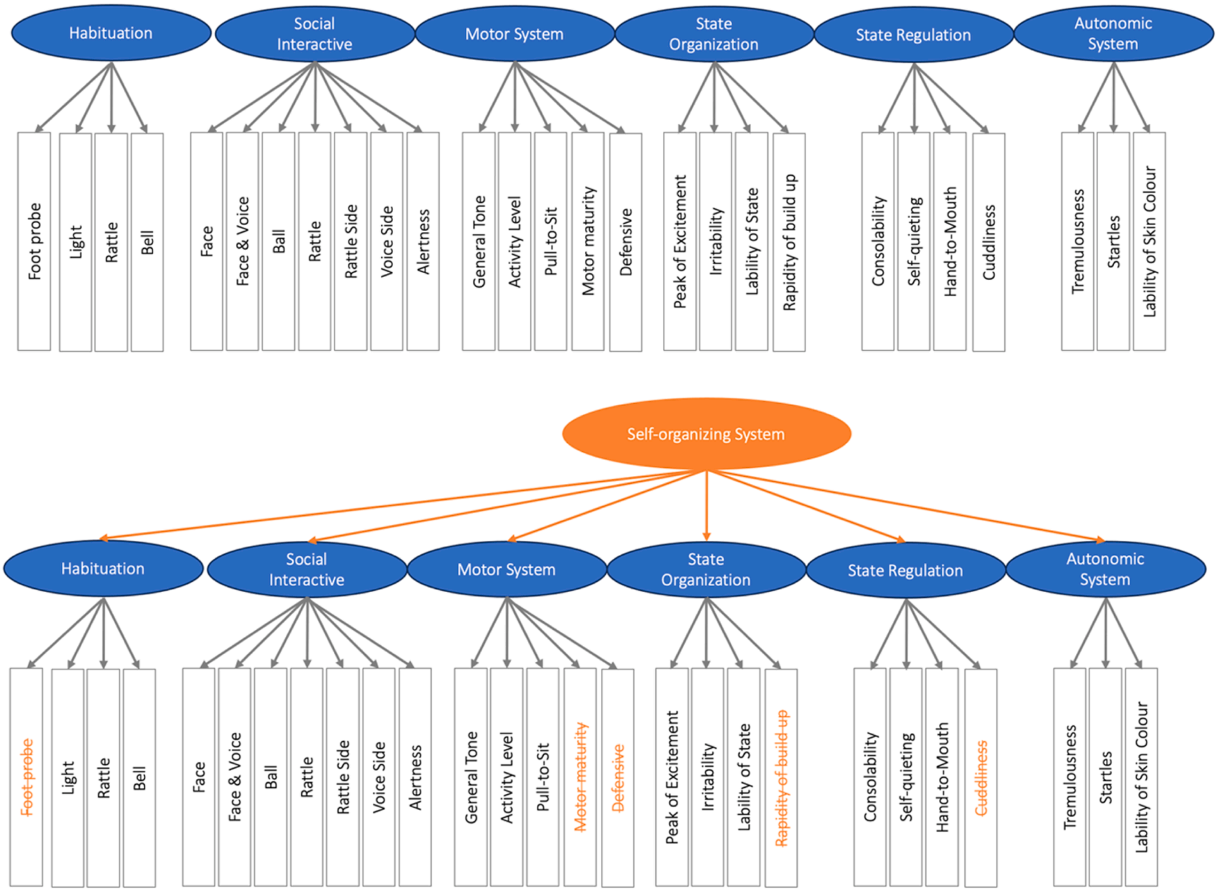

Fig. 2. Visualisation of the Lester model (top) corresponding to six original NBAS subscales; and modified model by Barbosa et al. (2018, bottom), including second-order Self-Organising System factor and deletion of highlighted items.

3.1.3. Procedure

The NBAS was administered when the child was 7–14 days of age. This age band was applied in both the UK and the Gambian cohorts, however it was informed by cultural practices in The Gambia which stipulate that the newborn should be strictly kept in the house with the mother for the first seven days of life (O'Neill et al., 2017). Assessments were conducted in participants' homes by two certified NBAS examiners, with one acting as an administrator and the second as an observer, in the presence of at least one caregiver. In The Gambia, the sessions were also videotaped by a village assistant or a third member of the research team. Observers took notes on an observation sheet and supervised the filming of the examination. Administrators filled in their own observation sheet at the end of the session. Both examiners completed the scoring form independently before completing a joint scoring form which was used as the final version. Any discrepancies were resolved by discussing observation sheets or consulting the session video.

3.1.4. Statistical analyses

Prior to analyses, items with mid-range or lowest optimal score were re-coded according to the NBAS manual (Brazelton & Nugent, 2011). Additionally, before recoding, Lability of State items scored as "null" due to infant remaining in one state throughout the examination (19 cases in the Gambian cohort) were assigned a raw score of 1 (corresponding to "1–2 state changes" in the manual) to aid data retention (Supplementary Table A). As per other factor analytic studies (e.g., Azuma et al., 1991), items examining newborn reflexes were not used in the current analyses. Statistical analyses were carried out in R, using the Lavaan 0.6.15 (Rosseel, 2012) package. To examine the factor structure of the NBAS we conducted CFA of two models: first, we fit the Lester factors (Fig. 2) corresponding to the subscales of habituation, orientation, motor system, state organisation, state regulation, and autonomic stability to our data separately per cohort.

Table 2
Item Level Scores Across the Gambia and the UK. Two-sample *t*-test significance shown for Benjamin-Hochberg corrected p values to control the false discovery rate (FDR) ($\alpha = 0.05$). *p < .05, * *p < .01, * **p < .001.

Item	Gambia (*N*=154)		UK (*N*=58)		
	N	Mean (SD)	N	Mean (SD)	t-statistic
Habituation					
Light	28	4.89 (3.49)	24	6.17 (2.76)	−1.47
Rattle	39	5.79 (3.22)	23	6.65 (2.31)	−1.22
Bell	27	6.78 (2.58)	19	6.53 (2.27)	0.35
Foot Probe	22	6.73 (2.25)	15	6.87 (2.56)	−0.17
Social Interactive/Orientation					
Animate Visual (Face)	150	4.61 (2.19)	52	4.83 (2.37)	−0.59
Animate Visual Auditory (Face & Voice)	146	4.83 (2.18)	47	5.43 (2.23)	−1.60
Inanimate Visual (Ball)	146	5.16 (2.55)	40	5.10 (2.48)	0.14
Inanimate Visual Auditory (Rattle)	148	5.98 (2.45)	34	5.41 (2.52)	1.19
Inanimate Auditory (Rattle Side)	147	5.67 (2.14)	35	5.11 (2.32)	1.28
Animate Auditory (Voice Side)	144	5.18 (1.91)	37	4.92 (2.43)	0.61
Alertness	151	5.30 (1.70)	52	5.38 (2.54)	−0.21
Motor System					
General Tone	154	5.84 (0.46)	58	5.74 (0.55)	1.19
Motor Maturity	153	7.24 (0.84)	57	5.95 (0.67)	11.56 * **
Pull-to-Sit	148	4.18 (2.08)	41	5.46 (1.82)	−3.89 * **
Defensive Movements	149	6.44 (2.13)	34	5.38 (2.20)	2.55
Activity Level	153	4.35 (0.52)	58	4.36 (0.83)	−0.08
State Organisation					
Peak of Excitement	154	4.41 (0.70)	58	3.83 (1.26)	3.33 * *
Rapidity of Build-up	153	2.24 (1.87)	57	2.91 (1.89)	−2.31
Irritability	154	4.80 (1.79)	58	4.76 (1.94)	0.14
Lability of States	154	2.53 (1.64)	57	3.56 (1.44)	−4.47 * **
State Regulation					
Cuddliness	148	4.43 (1.69)	50	5.50 (1.53)	−4.18 * **
Consolability	52	3.88 (1.68)	38	3.71 (1.96)	0.44
Self-Quieting	91	3.26 (1.90)	49	3.41 (1.99)	−0.42
Hand-to-Mouth	154	5.52 (1.92)	56	5.05 (2.50)	1.27
Autonomic System					
Tremulousness	154	7.97 (2.07)	58	7.67 (2.15)	0.92
Startles	142	7.60 (0.67)	32	7.50 (0.80)	0.65
Lability of Skin Colour	152	4.96 (0.23)	58	4.21 (0.93)	6.09 * **
Smiles	154	0.79 (1.22)	57	0.58 (1.32)	1.03
Supplementary Items					
Quality of Alertness	154	5.67 (1.78)	58	4.72 (2.38)	2.75
Cost of Attention	153	6.73 (1.18)	58	5.21 (1.82)	5.89 * **
Examiner Facilitation	154	5.97 (1.26)	58	4.91 (2.11)	3.60 * *
General Irritability	154	6.66 (1.41)	58	5.26 (2.23)	4.47 * **
Robustness and Endurance	153	6.46 (1.44)	58	4.71 (1.96)	6.18 * **
State Regulation	109	6.97 (0.75)	57	5.86 (2.13)	3.83 * *
Examiner's Emotional Response	153	7.62 (1.22)	58	7.38 (1.62)	1.03

We then examined the second-order modified factorial model introduced by Barbosa et al. (2018, Fig. 2), which in addition to the above factors included a second-order factor on self-organisation, as well as the deletion of five items from their corresponding subscales specifically: foot probe (habituation subscale), motor maturity and defensive manoeuvre (motor system subscale), rapidity of build-up (state organisation subscale) and cuddliness (state regulation subscale) as these were found to lead to poorer model fit in Barbosa et al.'s analyses. We then assessed the fit of both models. Model fit was considered satisfactory with a Comparative Fit Index (CFI) of > 0.90, a Tucker Lewis Index (TLI) of > 0.90 and a Root Mean Square Error of Approximation (RMSEA) of < .05 (Brown 2015; Hu & Bentler 1999). For the best fitting model, we extracted latent factor scores for subsequent analyses. We first examined associations between factor scores and individual items of the Supplementary subscale, to assess to what extent our factor model showed congruence with examiners' qualitative impressions of the infant's behaviour during the session. We then examined the perinatal factors of infant birthweight and gestational age, as well as maternal anxiety and anaemia and their association with NBAS factor scores to assess their relevance in predicting neonatal responses in context of rural Gambia, and to provide an insight into the potential mechanisms associated with early neonatal behaviour and mental health.

3.2. Results

3.2.1. Sample characteristics

For full details on the two study cohorts please see (Lloyd-Fox, McCann, Milosavljevic, Katus et al., 2024). In The Gambia, 204 families were recruited and eligible at birth. Of these, N = 154 participated in the NBAS, with data missing due to mothers travelling to other parts of the country for the birth and participants being missed within the permissible age bracket. Infants had a mean age of 12.4 days (SD=4.1 days). Maternal average parity was 4.43 (SD=2.37).

In the UK, 61 families were recruited and eligible at birth. Of these, 58 participated in the NBAS, with missing data being due to participants not being seen within the permissible age bracket. Infants on average were 12.2 days (SD=3.3 days). Maternal average parity was 1.19 (SD=0.4).

As discussed in more detail in (Lloyd-Fox, McCann, Milosavljevic, Katus et al., 2024), the target cohort sizes (The Gambia n = 200, UK n = 60,) were based on other measures implemented as part of this project, specifically previous infant fNIRS and EEG studies. These indicated that sample sizes from 20 (moderate effect size) to 42 (small effect size) were sufficient to determine regions of significant cortical brain activation in response to stimuli. The Gambian cohort was designed to be larger to allow within-cohort subgroup comparisons and individual differences analyses; for example, grouping by growth trajectories (mild, medium and severe markers of undernutrition) on the assumption that approximately 25–30 % of the cohort would be stunted (z-score of length-for-age < 2 standard deviations below the WHO reference) by two years of age (Nabwera et al., 2017).

3.2.2. NBAS Item Level Scores Across UK and Gambia

Significant between-site differences were discovered between infant scores for seven main items and six supplementary items (Table 2). At both sites, the items that depend on an infant being in a particular state, (e.g., asleep for the response decrement items or crying for consolability), had the largest proportion of missing data. Despite the UK site having fewer missing data than the Gambian site (approximately 65 % of UK infants had missing data across the response decrement items compared with 80 % of Gambian infants) the UK cohort was too small for robust model estimation. Hence, subsequent factor analytic approaches were performed solely on the Gambian cohort.

Item-level analyses showed pronounced differences in scoring between the UK and Gambian study sites (Table 1). Specifically, motor maturity, peak of excitement, and lability of skin colour were rated higher in the Gambian cohort, whereas higher scores were seen in the UK on pull-to-sit, lability of states, and cuddliness. Further, there were systematic differences on the supplementary

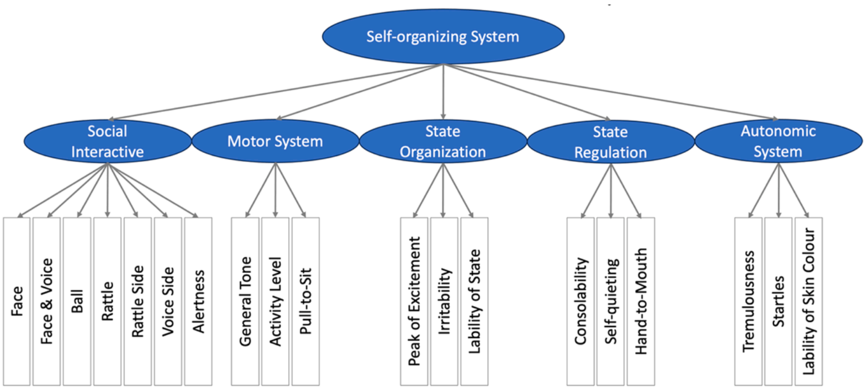

Fig. 3. Final model applied in current study: In line with Barbosa et al. (2018), a second- order Self-organising System factor was included, and items of foot probe (habituation subscale), motor maturity and defensive manoeuvre (motor system subscale), rapidity of build-up (State--Organisation subscale) and cuddliness (state regulation subscale) were omitted. Further, due to large proportions of missing data, the habituation subscale was omitted.

subscale, with five out of the seven items (cost of attention, examiner facilitation, general irritability, robustness and endurance, and state regulation), being rated higher in the Gambian cohort. In part, these differences underscore a qualitative observation made by examiners from each study sites which emphasise that infants in the Gambia required more stimulation to respond, were sleepier, and show fewer state changes throughout the session. Furthermore, some items, such as changes in skin colour, may be harder to assess under varying lighting conditions in participants' homes. While we anticipated that skin colour changes may be harder to observe on darker skin tones, examiners flagged primarily that a lack of electrical lighting in many participants' homes made skin colour changes more difficult to observe reliably.

3.2.3. Confirmatory Factor Analysis

Within the Gambian cohort, we conducted CFA using the original six Lester factors (Fig. 1). A good model fit was not achieved (CFI=.555, TLI=.494, RMSEA=.093). Fitting the Barbosa et al. (2018) model led to slight improvements in the model fit indices, though model fit overall was still poor (CFI=.677, TLI=.633, RMSEA=.080). Upon further inspection of the raw data as well as model-modification indices, it became apparent that large proportions of missing data within the habituation subscale were the primary cause for poor model fit. For this reason, the subscale was removed from analyses, and the CFA rerun. Using this approach, both the original Lester model (CFI=.917, TLI=.904, RMSEA=.036) as well as the Barbosa model (CFI=.958, TLI=.951, RMSEA=.028) yielded excellent fit indices. Due to providing the best fit as well as the conceptual importance of including the second-order Self-Organising System factor, the model by Barbosa et al. (2018) with omission of the habituation subscale was taken forward for further analyses (Fig. 3).

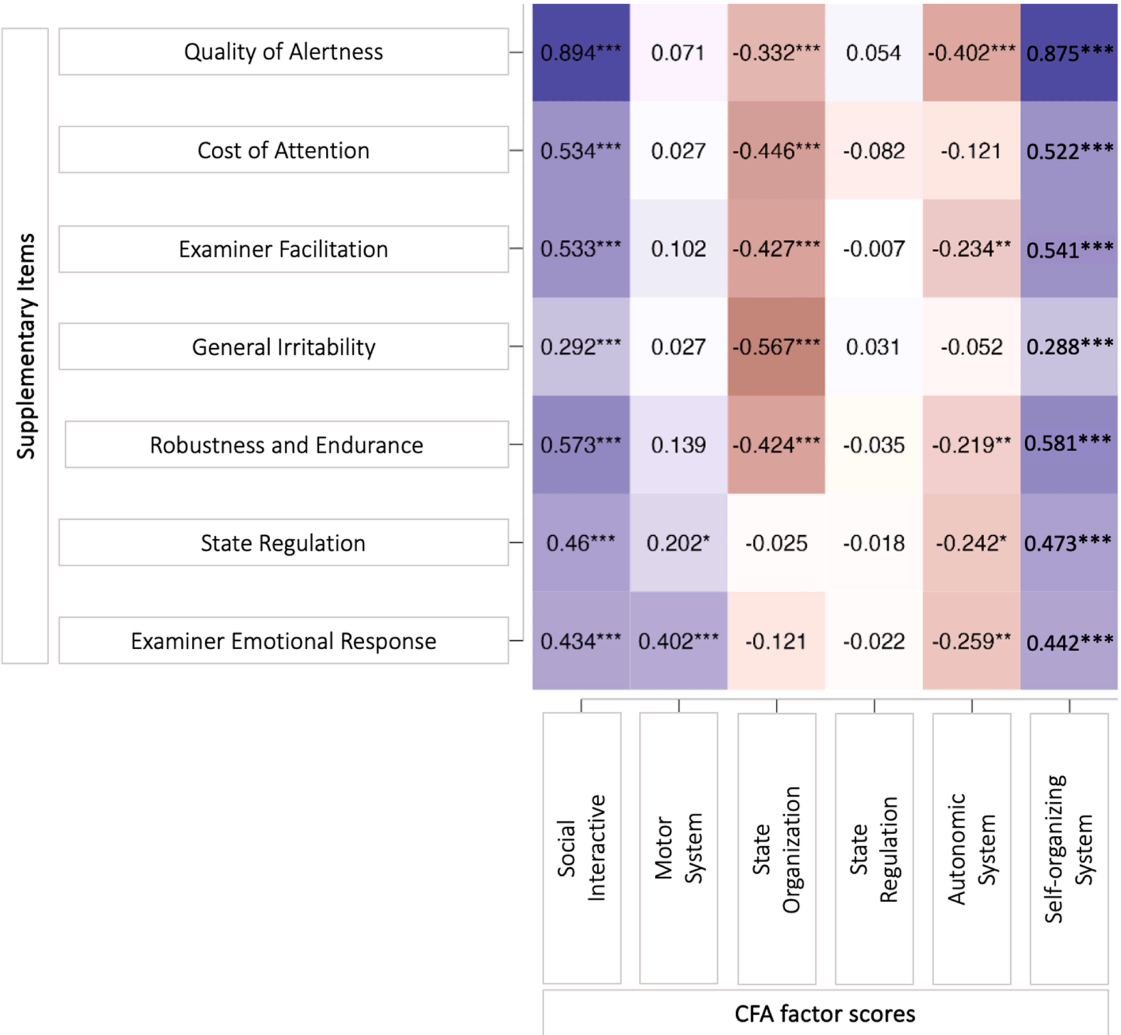

Fig. 4. Correlation matrix (two-tailed Pearson correlations) between the CFA factor scores and NBAS supplementary items (prior to re-coding, N = 154, FDR corrected), *p < .05, * *p < .01, * **p < .001.

Factor scores were extracted for the five factors mapping onto the corresponding subscales (Social Interactive, Motor System, State Organisation, State Regulation, Autonomic System) and for the second-order Self-Organising System factor.

3.2.4. Association of factor scores with Supplementary Items

To assess whether CFA factor scores showed agreement with the examiner's more qualitative impressions of the infant's behaviour during the NBAS session, we examined Pearson correlations between the factor scores and the NBAS supplementary items, which do not typically feature in factor analytic approaches of the scale. Results of the correlations (corrected for multiple comparisons via false-discovery rate correction (FDR, Benjamini & Hochberg, 1995)) are visualised in Fig. 4.

The Social Interactive and Self-Organising System factors were positively correlated with all items from the Supplementary sub-scale. An opposite pattern was observed for the Autonomic System and State-Organisation subscales, which showed a pattern of negative correlations with the Supplementary items, due to being negatively scored (i.e., higher scores indicating poorer performance). The factor scores for the Motor System and State Regulation subscales showed few associations with the Supplementary items. This may, in part, be accounted for by the fact that despite suggesting a close link to specific NBAS subscales by name (e.g., State Regulation Supplementary Items and State Regulation subscale), what is being measured as part of the supplementary scale is an independent examiner impression rather than an additional item belonging to the specific scale.

3.2.5. Associations with maternal and infant perinatal factors

Descriptive statistics for infant (birthweight, gestational age) and maternal (antenatal anxiety, anaemia) factors are summarised in Table 3.

To assess associations between infant (birthweight, gestational age) and maternal (antenatal anxiety, anaemia) factors and NBAS factor scores, we examined Pearson correlations between the NBAS factor scores (Social Interactive, Motor System, State-Organisation, State Regulation, Autonomic System, Self-Organising System) and infant birthweight, gestational age at birth and maternal antenatal anxiety. FDR-corrected results can be seen in Fig. 5.

Some associations between perinatal risk factors and NBAS factor scores were observed: most consistently, these were seen in relation to the Social Interactive factor and the Self-Organising System Factor. Associations were found between the Social-Interactive factor and pregnancy anxiety, gestational age at birth and birthweight. For the Self-Organising system, associations were seen with pregnancy anxiety, gestational age at birth and birthweight, with birthweight not surviving correction. Pregnancy anxiety was negatively associated with the two NBAS factors, with higher anxiety being associated with lower NBAS scores, whereas positive associations were seen for all other significant correlations.

Further, factor scores were compared via independent t-tests for anaemic vs non-anaemic mothers. No significant differences were seen for any NBAS factor and maternal anaemia (p all $>.174$).

4. Discussion

While inclusion of a more diverse set of populations in neonatal developmental research was proposed over 20 years ago (Sander, 2000), assessment measures of newborn behaviour implemented in low-and middle-income settings remains scarce. Our study contributes to bridging this gap by examining the utility of the NBAS in rural Gambia, West Africa.

Our pilot study yielded valuable insights into the acceptability and feasibility of implementing the NBAS in this rural sub-Saharan African context. Most mothers, fathers, and grandparents expressed positive or neutral feedback regarding the session, suggesting overall acceptability and underscoring the cross-cultural utility highlighted in prior research. Findings emphasized the importance of clear communication, pre-session explanation, and sensitivity to parental comfort. Analysis of item-specific feedback provided further insight into the adaptation process. Namely, habituation items (particularly to light), raised concerns among some participants and were thus introduced with special care. Additionally, challenges related to specific motor and reflex items, including pull-to-sit and placing manoeuvres, necessitated adjustments and, in some cases, omission where the home environment did not offer a suitable set-up. One recurring theme throughout the pilot study was the recognition that parents permitted certain actions only as being part of a research study, underscoring examiners' ethical responsibility toward caregivers' degree of comfort while conducting assessments.

Building upon the insights from the pilot study, the BRIGHT project extended the investigation to a larger cohort in The Gambia and the UK. While a thorough evaluation of the NBAS' psychometric properties in the UK cohort was precluded by limitations in sample size, item level comparisons of raw scores provide some insight into specific differences across the sites: we saw some systematic differences in items primarily relating to infants' propensity to change and maintain different states, as well as the degree to which they required an examiner's support to be able to show a specific behaviour, with Gambian infants overall showing fewer state changes and

Table 3
Descriptive statistics of maternal and infant factors.

	Gambia (N=154)	
	N	**Mean (SD)**
Birthweight	129	3.061 (.340)
Gestational age at birth	152	39.799 (1.803)
Antenatal Anxiety	132	12.5 (3.783)
Maternal Haemoglobin	131	10.951 (1.171) (18 % anaemic)

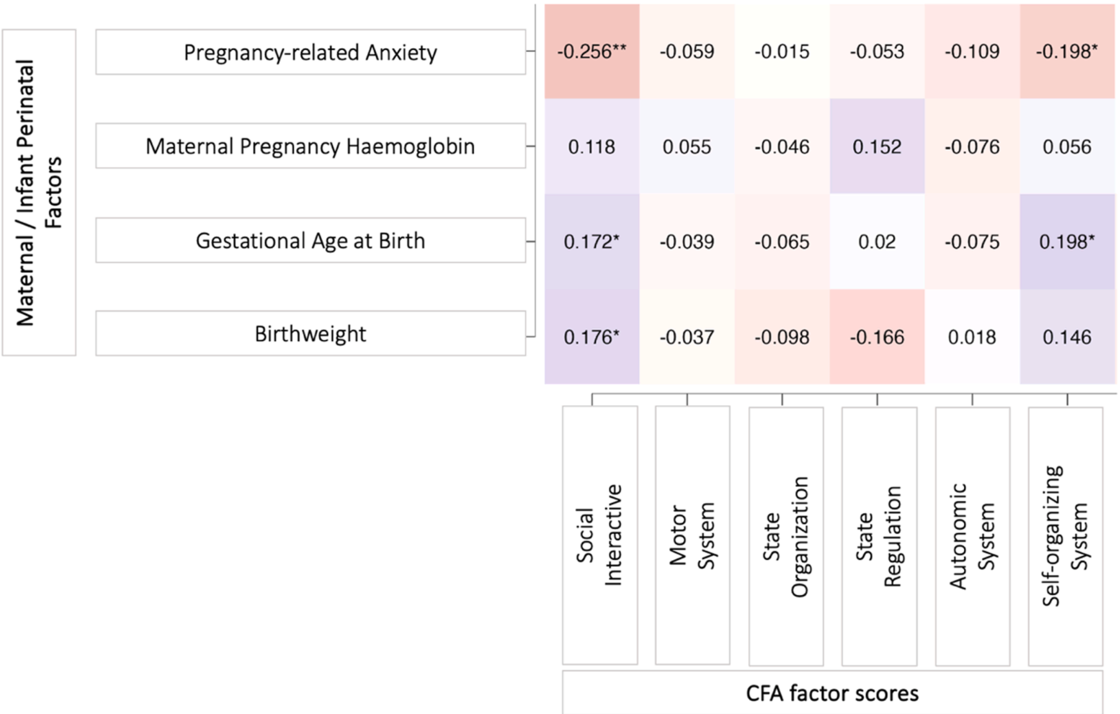

Fig. 5. Correlation matrix (two-tailed Pearson correlations) between the NBAS CFA factor scores and maternal/infant perinatal factors (N = />129, FDR corrected), *p < .05, * *p < .01, * **p < .001.

a greater degree of examiner facilitation throughout the session. It needs to be noted that without further psychometric evaluation of these differences it is unclear whether they represent a true mean-level difference or merely a systematic difference in how administrators made use of the scoring criteria. While administration was standardised to a high degree, such a confounding effect cannot be ruled out.

Factor analysis within the larger Gambian cohort aimed at affirming the factor structure of the NBAS. The adaptation of the model by Barbosa et al. (2018), including its added second-order Self-organising Factor), with the omission of the habituation subscale, necessitated by large proportions of missing data, led to an excellent model fit. This highlights the by-now recurrent finding that, other than the individual subscale-scores, an overarching construct of an infant's ability to self-regulate may be needed to capture all variance seen in NBAS scores (Barbosa et al., 2018, McCollam et al., 1997). Validating this partially data-driven approach, we found consistent positive associations between the NBAS factor scores for social orientation and the second-order Self-Organising factor, as well as a reverse pattern for the negatively coded factors of State Organisation and the Autonomic System, and examiners' more qualitative impressions of the session via NBAS supplementary items. However, no such associations were found for the NBAS factors of state regulation and motor system, which cautions against treating the supplementary items as mere additional items for these subscales. Finally, we assessed the relationship between maternal and infant perinatal factors and the NBAS scores. The most consistent associations were found with the Social Interactive factor (specifically with pregnancy anxiety, gestational age at birth and birthweight), as well as the Self-Organising System factor and pregnancy anxiety and gestational age at birth. With most prior work in the domain of maternal mental health focussing on links with depression, our finding adds to the limited literature of how anxiety relates to neonatal outcomes that warrants further investigation. Regarding birthweight, we found that higher birthweight was associated with Social Interactive factor scores; however, the associations with the Self-Organising system factor did not survive statistical correction.

These findings are in line with prior research, highlighting the impact of poor maternal mental health on NBAS/NBO scores, with effects for the social-interactive cluster being most reported (Gressier et al., 2020; Osborne et al., 2022, Zhang et al., 2020). Since only a minority of studies to date include a second-order factor, our association with such an outcome measure has less precedence in published literature but does warrant further investigation. While preterm birth and small size for gestational age have been highlighted as risk factors for poorer performance on the NBAS (e.g., Greene, Fox, Lewis, 1983, Alvarez-Garcia et al., 2015, Figueras et al., 2009), our data only partly corroborates this, by showing strongest links between gestational age at birth and birthweight with the social interactive factor score.

We did not, however, find associations between the remaining NBAS factors (motor system, state organisation, state regulation and autonomic system) and maternal anxiety, gestational age, birthweight or maternal anaemia. Indeed, contrary to prior research indicating lower NBAS scores in infants of iron-deficient mothers (Hernández-Martínez et al., 2011), maternal anaemia did not show an association with any NBAS factor scores in our study. This may in part be due to relative protective mechanisms to maximise the foetal endowment of iron, which ensure that, even in contexts of severe iron deficiency, a foetal supply is maintained meaning iron deficiency and its associated neurodevelopmental consequences may not manifest until later on (Domello et al., 2014).

In summary, our findings highlight a pattern by which those factors that do show an association with NBAS outcomes (specifically pregnancy anxiety, gestational age at birth and, to an extent, birthweight) have their primary impacts on the social-interactive responses, as well as the overall self-regulatory abilities of the infant. These results warrant further follow up of maternal and infant risk factors across associated developmental delays beyond the newborn period, as well as a closer examination of the long-term predictive utility of neonatal social and self-regulatory abilities across infancy and early childhood.

4.1. Strengths and Limitations

While our study contributes to the understanding of NBAS applicability in diverse settings, certain limitations should be acknowledged. The exclusion of the UK cohort from follow-on analyses due to missing data emphasizes the challenges associated with obtaining robust results from data reduction techniques. Further, missing data on the habituation subscale of the NBAS even in the Gambian sample highlights that where indices on specific subscales are of particular relevance, strategies of optimising testing protocols so that these responses can be obtained more easily should be implemented (e.g., where habituation is a key outcome, infants wherever possible should be seen at a time when they are more likely to be asleep at the beginning of the session).

There also are a range of notable strengths: first, assessments in Phase 2 were carried out by trained local researchers highly familiar with the context and therefore able to relate to families participating in the study. This approach necessitated a remote training protocol, which we were able to successfully implement, and this could be applied for other research desiring local researchers to utilise the NBAS in settings without a training centre. Lastly, our partly data-driven approach made fewer assumptions about the generalisability of newborn behaviour across settings, and resulting models may represent a useful data reduction method for future cross-cultural work.

4.2. Conclusion

In conclusion, our study responds to the need of expanding research on neonatal behaviour to encompass diverse global populations. The successful adaptation and implementation of the NBAS in rural Gambia underscores the potential for such standardised assessments to contribute to this gap in research. Our findings demonstrate that the NBAS can be introduced sensitively across diverse settings, while also highlighting the scales' utility in making research on newborn behaviour and infant mental health more inclusive and globally relevant. The insights gained pave the way for future studies seeking to understand the interplay between cultural contexts, perinatal factors, and early neurobehavioural development.

Funding information

Medical Research Council UK and the UK Department for International Development (DfID) under the MRC/DfID Concordant agreement., Grant/Award Number: MC-A760-5QX00; Bill and Melinda Gates Foundation, Grant/Award Number: OPP1127625.

Conflict of interest disclosure

The authors declare no conflict of interest.

Ethics approval statement

For the pilot work, ethical approval was granted by the joint Gambia Government – MRC Unit The Gambia Ethics Committee for field observation, caregiver interviews and pilot NBAS sessions (project title: A pilot study for the implementation of the Neonatal Behavioural Assessment Scale (NBAS) in The Gambia", SCC number 1413v2). Additionally, full approval was granted by the Warwick Biomedical & Scientific Research Ethics Committee (BSREC) for the NBAS pilot sessions (project title: The Neonatal Behavioral Assessment Scale (NBAS) for supporting caregivers and newborns in Keneba, the Republic of The Gambia: development, piloting and feasibility, REGO-2014–1300 AM01). For the main study in The Gambia, ethical approval was obtained from the joint Gambia Government – MRC Unit The Gambia Ethics Committee (project title 'Developing brain function for age curves from birth using novel biomarkers of neurocognitive function', SCC number 1451v2). In the UK, ethical approval was granted by the NHS Health Research Authority (reference 15/EE/0202, project 178682).

Permission to reproduce material from other sources

Not applicable.

CRediT authorship contribution statement

Sophie E. Moore: Writing – review & editing, Supervision, Resources, Funding acquisition. **Clare E. Elwell:** Writing – review & editing, Supervision, Resources, Funding acquisition. **Sarah Lloyd-Fox:** Writing – review & editing, Supervision, Funding acquisition, Conceptualization. **Mustapha Minteh:** Investigation. **Samantha McCann:** Writing – review & editing, Project administration, Data curation. **Jane Barlow:** Supervision, Resources, Conceptualization. **Malang Jammeh:** Investigation. **Fabakary Njai:** Validation, Project administration, Investigation. **Tijan Fadera:** Validation, Project administration, Investigation. **Bosiljka Milosavljevic:** Writing – review & editing, Validation, Data curation. **Marta Perapoch Amadó:** Writing – review & editing, Investigation, Data curation. **Laura Katus:** Writing – review & editing, Writing – original draft, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Christine Torrance:** Writing – review & editing, Writing – original draft, Supervision, Project administration, Investigation, Formal analysis, Data curation, Conceptualization. **Maria Rozhko:** Writing – review & editing, Writing – original draft, Project administration, Investigation, Formal analysis, Data curation, Conceptualization.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.infbeh.2024.102017.

Data availability

Data will be made available on request.

References

Achenbach, T. M., & Rescorla, L. A. (2000). Child behavior checklist for ages 1 1/2-5. *Reporter, 10*, 20.
Aliabadi, F., & Askary, R. K. (2013). Effects of Tactile–kinesthetic Stimulation on low birth Weight neonates. *Iranian journal of pediatrics, 23*(3), 289.
Alvarez-Garcia, A., Fornieles-Deu, A., Costas-Moragas, C., & Botet-Mussons, F. (2015). Neurobehavioral conditions and effects of gender, weight and severity in preterm infants according to the Neonatal Behavioral Assessment Scale. *Anales Délelőtt Psicología/Annals of Psychology, 31*(3), 818–824.
Ayala, A., Christensson, K., Christensson, E., Cavada, G., Erlandsson, K., & Velandia, M. (2021). Newborn infants who received skin-to-skin contact with fathers after Caesarean sections showed stable physiological patterns. *Acta paediatrica, 110*(5), 1461–1467.
Ayesa-Arriola, R., Castro Quintas, A., Ortiz-García de la Foz, V., Miguel Corredera, M., San Martín González, N., Murillo-García, N. , … de Las Cuevas-Terán, I. (2023). Exploring the impact of COVID-19 on newborn neurodevelopment: a pilot study. *Scientific Reports, 13*(1), Article 2983.
Azuma, S. D., Malee, K. M., Kavanagh, J. A., & Deddish, R. B. (1991). Confirmatory factor analysis with preterm NBAS data: A comparison of four data reduction models. *Infant Behavior and Development, 14*(2), 209–225.
Barbosa, M., Moreira, J., Tronick, E., Beeghly, M., & Fuertes, M. (2018). Neonatal Behavioral Assessment Scale (NBAS): Confirmatory factor analysis of the six behavioral clusters. *Early Human Development, 124*, 1–6.
Belot, R. A., Bouteloup, M., Mariage, A., Candilis-Huisman, D., Mottet, N., & Mellier, D. (2021). Relevance of Neonatal Behavior Assessment Scale for infants with somatic disorders: comparison on one matched group of control. *Frontiers in Pediatrics, 8*, Article 506384.
Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical society: Series B (Methodological), 57*(1), 289–300.
Brazelton, T. B., & Nugent, J. K. (2011). *Neonatal Behavioral Assessment Scale* (4th ed.). London: MacKeith Press.
Briggs-Gowan, M. J., Carter, A. S., Irwin, J. R., Wachtel, K., & Cicchetti, D. V. (2004). The Brief Infant-Toddler Social and Emotional Assessment: screening for social-emotional problems and delays in competence. *Journal of Pediatric Psychology, 29*(2), 143–155.
Canals, J., Esparó, G., & Fernández-Ballart, J. D. (2006). Neonatal behaviour characteristics and psychological problems at 6 years. *Acta Paediatrica, 95*(11), 1412–1417.
Canals, J., Hernández-Martínez, C., Esparó, G., & Fernández-Ballart, J. (2011). Neonatal Behavioral Assessment Scale as a predictor of cognitive development and IQ in full-term infants: a 6-year longitudinal study. *Acta Paediatrica, 100*(10), 1331–1337.
Carter, A. S., Briggs-Gowan, M. J., Jones, S. M., & Little, T. D. (2003). The infant–toddler social and emotional assessment (ITSEA): Factor structure, reliability, and validity. *Journal of Abnormal Child Psychology, 31*, 495–514.
Cecil, C. A., Pickles, A., Hill, J., & Sharp, H. (2017). The joint contributions of maternal depression and infant negative emotionality to maternal sensitivity to infant distress.
Chen, X. K., Chen, S. H., Lv, G. R., You, J. H., & Chen, Z. K. (2017). Clinical value of transfontanellar ultrasonography for neonatal insular development. *Clinical and Experimental Obstetrics & Gynecology, 44*(5), 758–763.
Clinton, J., Feller, A. F., & Williams, R. C. (2016). The importance of infant mental health. *Paediatrics Child Health, 21*(5), 239–241.
DeGangi, G.A., Poisson, S., Sickel, R., & Wiener, A. (1995). Infant. Toddler Symptom Checklist: A screening tool for parents Psychological Corp.
Eishima, K. (1992). The effects of obstetric conditions on neonatal behaviour in Japanese infants. *Early Human Development, 28*(3), 253–263.
Feldman, R. (2015). Mutual influences between child emotion regulation and parent–child reciprocity support development across the first 10 years of life: Implications for developmental psychopathology. *Development and psychopathology, 27*(4pt1), 1007–1023.
Feldman, R., & Eidelman, A. I. (2006). Neonatal state organization, neuromaturation, mother-infant interaction, and cognitive development in small-for-gestational-age premature infants. *Pediatrics, 118*(3), e869–e878.
Figueras, F., Oros, D., Cruz-Martinez, R., Padilla, N., Hernandez-Andrade, E., Botet, F., … Gratacos, E. (2009). Neurobehavior in term, small-for-gestational age infants with normal placental function. *Pediatrics, 124*(5), e934–e941.
Froggatt, S., Covey, J., & Reissland, N. (2020). Infant neurobehavioural consequences of prenatal cigarette exposure: A systematic review and meta-analysis. *Acta Paediatrica, 109*(6), 1112–1124.
Galbally, M., Stein, A., Hoegfeldt, C. A., & van IJzendoorn, M. (2020). From attachment to mental health and back. *The Lancet Psychiatry, 7*(10), 832–834.
Greene, J. G., Fox, N. A., & Lewis, M. (1983). The relationship between neonatal characteristics and three-month mother-infant interaction in high-risk infants. *Child Development*, 1286–1296.
Gressier, F., Letranchant, A., Glatigny-Dallay, E., Falissard, B., & Sutter-Dallay, A. L. (2020). Negative impact of maternal antenatal depressive symptoms on neonate's behavioral characteristics. *European Child Adolescent Psychiatry, 29*(4), 515–526.
Groh, A. M., Fearon, R. P., van IJzendoorn, M. H., Bakermans-Kranenburg, M. J., & Roisman, G. I. (2017). Attachment in the early life course: Meta-analytic evidence for its role in socioemotional development. *Child Development Perspectives, 11*(1), 70–76.
Grossmann, K., Grossmann, K. E., Spangler, G., Suess, G., & Unzner, L. (1985). Maternal sensitivity and newborns' orientation responses as related to quality of attachment in northern Germany. *Monographs of the society for Research in Child Development*, 233–256.
Gunning, M., Halligan, S. L., & Murray, L. (2013). Contributions of maternal and infant factors to infant responding to the Still Face paradigm: A longitudinal study. *Infant Behavior and Development, 36*(3), 319–328.

Hennig, B. J., Unger, S. A., Dondeh, B. L., Hassan, J., Hawkesworth, S., Jarjou, L., & Fulford, A. J. (2017). Cohort profile: the Kiang West Longitudinal Population Study (KWLPS)—a platform for integrated research and health care provision in rural Gambia. *International Journal of Epidemiology, 46*(2). e13-e13.

Katus, L., Crespo-Llado, M. M., Milosavljevic, B., Saidykhan, M., Njie, O., Fadera, T., … BRIGHT Project Team. (2024). It takes a village: Caregiver diversity and language contingency in the UK and rural Gambia. *Infant Behavior and Development, 74*, Article 101913.

Katus, L., Hayes, N. J., Mason, L., Blasi, A., McCann, S., Darboe, M. K., … Elwell, C. E. (2019). Implementing neuroimaging and eye tracking methods to assess neurocognitive development of young infants in low-and middle-income countries. *Gates Open Research, 3*.

Katus, L., Mason, L., Milosavljevic, B., McCann, S., Rozhko, M., Moore, S. E., … Prentice, A. (2020). ERP markers are associated with neurodevelopmental outcomes in 1–5 month old infants in rural Africa and the UK. *NeuroImage, 210*, Article 116591.

Kooistra, L., Crawford, S., van Baar, A. L., Brouwers, E. P., & Pop, V. J. (2006). Neonatal effects of maternal hypothyroxinemia during early pregnancy. *Pediatrics, 117*(1), 161–167.

Lloyd-Fox, S., McCann, S., Milosavljevic, B., Katus, L., Blasi, A., Bulgarelli, C., … BRIGHT Project Team. (2024). The Brain Imaging for Global Health (BRIGHT) project: longitudinal cohort study protocol. *Gates Open Research, 7*, 126.

Lloyd-Fox, S., Blasi, A., McCann, S., Rozhko, M., Katus, L., Mason, L., … Bright Project Team. (2019). Habituation and novelty detection fNIRS brain responses in 5- and 8-month-old infants: The Gambia and UK. *Developmental Science, 22*(5), Article e12817.

Loo, K. K., Ohgi, S., Zhu, H., Howard, J., & Chen, L. (2005). Cross-cultural comparison of the neurobehavioral characteristics of Chinese and Japanese neonates. *Pediatrics International, 47*(4), 446–451.

Lowinger, S. (1999). Infant irritability and early mother–infant reciprocity patterns. Infant and Child Development. *An International Journal of Research and Practice, 8*(2), 71–84.

Malak, R., Fechner, B., Sikorska, D., Rosołek, M., Mojs, E., Samborski, W., & Baum, E. (2021). Application of the Neonatal Behavioral Assessment Scale to Evaluate the Neurobehavior of Preterm Neonates. *Brain Sciences, 11*(10), 1285.

Malak, R., Wiecheć, K., Fechner, B., Szczapa, T., Kasperkowicz, J., Matthews-Kozanecka, M. , … Mojs, E. (2022). The influence of parent education on the neurobehavior and sucking reflexes of very preterm infants. *Brain Sciences, 12*(7), 840.

McCann, S., Mason, L., Milosavljevic, B., Mbye, E., Touray, E., Colley, A., & BRIGHT Study Team. (2023). Iron status in early infancy is associated with trajectories of cognitive development up to pre-school age in rural Gambia. *PLOS Global Public Health, 3*(11), Article e0002531.

Medoff-Cooper, B., & Ratcliffe, S. J. (2005). Development of preterm infants: feeding behaviors and brazelton neonatal behavioral assessment scale at 40 and 44 weeks' postconceptional age. *Advances in Nursing Science, 28*(4), 356–363.

O'Neill, S., Clarke, E., & Grietens, K. P. (2017). How to protect your new-born from neonatal death: infant feeding and medical practices in the Gambia (January). In *In Women's Studies International Forum, 60* pp. 136–143). Pergamon (January).

Ohgi, S., Takahashi, T., Nugent, J. K., Arisawa, K., & Akiyama, T. (2023). Neonatal behavioral characteristics and later behavioral problems. *Clinical pediatrics, 42*(8), 679–686.

Osborne, S., Biaggi, A., Hazelgrove, K., Du Preez, A., Nikkheslat, N., Sethna, V., & Pariante, C. M. (2022). Increased maternal inflammation and poorer infant neurobehavioural competencies in women with a history of major depressive disorder from the psychiatry research and motherhood–Depression (PRAM-D) study. *Brain, Behavior, and Immunity, 99*, 223–230.

Puhlmann, L. M., Derome, M., Morosan, L., Kilicel, D., Vrtička, P., & Debbané, M. (2021). Longitudinal associations between self-reported attachment dimensions and neurostructural development from adolescence to early adulthood. *Attachment Human Development*, 1–19.

Purpura, G., Fumagalli, S., Nacinovich, R., Riva, A., Ornaghi, S., Serafini, M., & Nespoli, A. (2023). Effects of social and sensory deprivation in newborns: A lesson from the Covid-19 experience. *Early Human Development, 185*, Article 105853.

Rencken, G., Govender, P., & Uys, C. J. (2022). Neurobehavioural challenges experienced by HIV exposed infants: a study in South Africa. *BMC pediatrics, 22*(1), 479.

Richter, L. M. (2003). Poverty, underdevelopment and infant mental health. *Journal of Paediatrics and Child Health, 39*(4), 243–248.

Sagiv, S. K., Nugent, J. K., Brazelton, T. B., Choi, A. L., Tolbert, P. E., Altshul, L. M., & Korrick, S. A. (2008). Prenatal organochlorine exposure and measures of behavior in infancy using the Neonatal Behavioral Assessment Scale (NBAS). *Environmental Health Perspectives, 116*(5), 666–673.

Sameroff, A., Seifer, R., & McDonough, S.C. (2004). Contextual contributors to the Assessment of Infant Mental Health Handbook of Infant, Toddler, and Preschool Mental Health Assessment6176.

Sander, L. W. (2000). Where are we going in the field of infant mental health? *Infant Mental Health Journal, 21*(1-2), 5–20.

Shoaff, J. R., Nugent, K., Brazelton, T. B., & Korrick, S. A. (2021). Early infant behavioural correlates of social skills in adolescents. *Paediatric and Perinatal Epidemiology, 35*(2), 247–256.

Soler-Limón, K. M., Rivera-González, R., Velázquez-Moctezuma, J., González-Camarena, R., Ontiveros-Mendoza, E., & Sánchez-Pérez, C. (2019). Neurobehavior of mexican newborns in conditions of low perinatal risk. *Acta Pediátrica Délelőtt Megyeéxico, 40*(3), 113–121.

Squires, J., Bricker, D., & Twombly, E. (2002). *Ages and stages questionnaires: Social-emotional*. Baltimore: Brookes.

Stengel, T. J. (1981). The neonatal behavioral assessment scale: description, clinical uses, and research implications. *Physical & Occupational Therapy In Pediatrics, 1*(1), 39–57.

Stewart, P., Reihman, J., Lonky, E., Darvill, T., & Pagano, J. (2000). Prenatal PCB exposure and neonatal behavioral assessment scale (NBAS) performance. *Neurotoxicology and teratology, 22*(1), 21–29.

Super, C. M., & Harkness, S. (2020a). Culture and the perceived organization of newborn behavior: A comparative study in Kenya and the United States. *New Directions for Child and Adolescent Development, 2020*(172), 11–24.

Super, C. M., & Harkness, S. (2020b). Culture and the perceived organization of newborn behavior: A comparative study in Kenya and the United States. *New Directions for Child and Adolescent Development, 2020*(172), 11–24.

Szaniecki, E., & Barnes, J. (2016). Measurement issues: Measures of infant mental health. *Child and Adolescent Mental Health, 21*(1), 64–74.

Tabaczyńska, A., Malak, R., Fechner, B., Mojs, E., Samborski, W., & Baum, E. (2021). The Comparison of Three Assessment Scales in a Neonatal Unit in Poland. *Studies in Logic, Grammar and Rhetoric, 66*(3), 635–646.

Tirosh, E., Hard, J., Abadi, J., Berger, A., & Cohen, A. (1992). Relationship between neonatal behavior and subsequent temperament. *Acta Paediatrica, 81*(10), 829–831.

Tronick, E., Wise, S., Als, H., Adamson, L., Scanlon, J., & Brazelton, T. B. (1976). Regional obstetric anesthesia and newborn behavior: Effect over the first ten days of life. *Pediatrics, 58*(1), 94–100.

Wang, S., Ding, C., Dou, C., Zhu, Z., Zhang, D., Yi, Q., … Li, H. (2022). Associations between maternal prenatal depression and neonatal behavior and brain function–evidence from the functional near-infrared spectroscopy. *Psychoneuroendocrinology, 146*, Article 105896.

Wolf, M. J., Koldewijn, K., Beelen, A., Smit, B., Hedlund, R., & De Groot, I. J. M. (2002). Neurobehavioral and developmental profile of very low birthweight preterm infants in early infancy. *Acta Paediatrica, 91*(8), 930–938.

Zeanah, C.H. (Ed.). (2018). Handbook of infant mental health. Guilford Publications.

Zhang, H., Liu, S., Si, Y., Zhang, S., Tian, Y., Liu, Y., & Zhu, Z. (2019). Natural sunlight plus vitamin D supplementation ameliorate delayed early motor development in newborn infants from maternal perinatal depression. *Journal of Affective Disorders, 257*, 241–249.

Zhang, H., Shao, S., Su, Q., Yao, D., Sun, H., Ding, D., … Li, H. (2018). Involvement of prolactin in newborn infant irritability following maternal perinatal anxiety symptoms. *Journal of Affective Disorders, 238*, 526–533.