

A Smartphone App Effectively Facilitates Mothers' Mind-Mindedness: A Randomized Controlled Trial

Fionnuala Larkin, Janine Oostenbroek, Yujin Lee, Emily Hayward, Amy Fernandez,
Ying Wang, Alex Mitchell, Lydia Y. Li & Elizabeth Meins

Child Development, in press

1 Abstract

The efficacy of a smartphone app intervention (BabyMind©) in facilitating mind-mindedness was investigated in a randomized controlled trial, assigning mothers and their 6-month-olds ($N=152$; 72 girls, 146 White) to intervention or active control conditions. Mothers who had received the BabyMind© app intervention scored higher for appropriate ($d = .61$, 95% CI .28, .94) and lower for non-attuned ($d = -.55$, 95% CI -.92, -.18) mind-related comments at follow-up (age 12 months), compared with their control group counterparts. Adjusting for missing data did not alter this pattern of findings. Mothers' baseline parental reflective functioning did not moderate these relations. Results are discussed in terms of the benefits of early intervention and exploring the efficacy of the app in more diverse populations.

2 Introduction

Research over the last 25 years has shown that caregivers' mind-mindedness (Meins, 1997) predicts various positive aspects of child development (see Aldrich, Chen, & Alfieri, 2021; and McMahon & Bernier, 2017, for reviews). Mind-mindedness is defined as the caregiver's tendency to treat the young child as an individual with a mind of their own, and is operationalized in infancy in terms of caregivers' comments about their infants' thoughts and feelings. Mind-minded caregivers tend to comment appropriately on the infant's internal states (appropriate mind-related comments) rather than misinterpreting what the infant is thinking or feeling (non-attuned mind-related comments) (Meins, Fernyhough, Fradley, & Tuckey, 2001; Meins et al., 2012). Higher levels of mind-mindedness are associated with secure attachment (e.g., Lundy, 2003; Meins, Bureau, & Fernyhough, 2018; Meins et al., 2001, 2012), and children's own understanding of other minds in the preschool years (e.g., Centifanti, Meins, & Fernyhough, 2016; Fishburn, Meins, Fernyhough, Centifanti, & Larkin, 2022; Kirk et al., 2015; Laranjo, Bernier, Meins, & Carlson, 2010, 2014). In children from socially and economically disadvantaged backgrounds, mind-mindedness in the first year of life is reported to predict fewer behavioral difficulties in the preschool years and the first year of school (Meins, Centifanti, Fernyhough, & Fishburn, 2013) and higher attainment in standardized school assessments at ages 7 and 11 (Meins, Fernyhough, & Centifanti, 2019).

2.1 Interventions to Facilitate Mind-Mindedness

Due to the positive associations in the extant literature between mind-mindedness and children's outcomes, researchers have devised intervention procedures in order to attempt to facilitate caregivers' mind-mindedness. The first of these interventions used video-feedback to help caregivers to comment more appropriately on their children's internal states. Colonnese et al.'s (2012) intervention focused on adoptive parents and their children, and was associated with a decrease in insecure attachment. Schacht et al. (2017) used a video-feedback mind-mindedness intervention with their sample of mothers hospitalized for severe mental illness. They reported that, compared to mothers who had received video-feedback targeted at encouraging them to talk to their infants and follow their gaze, those who had received the mind-mindedness intervention showed a

significant decrease in non-attuned mind-related comments and a non-significant increase in appropriate comments from admission to discharge. Moreover, the mothers who had received the intervention were not significantly different from psychologically well controls in terms of either appropriate or non-attuned comments at discharge. At follow-up in the second year of life, mothers who had received the intervention were also more likely to have secure attachment relationships with their infants compared with the comparison group of mothers who had also been hospitalized for severe mental illness.

Larkin, Oostenbroek, Lee, Hayward, & Meins (2019) investigated whether it was possible to facilitate mind-mindedness using an intervention that was less labor-intensive than the individually-delivered video-feedback procedures that had proved effective in previous research. They developed a smartphone app that provided caregivers with psychoeducation on early infant development and sent a daily alert prompting the caregiver to consider what was “on their infant’s mind”. In response to these alerts, the caregiver could post photographs or video clips to indicate how the infant was feeling or what they were thinking, with the research team providing feedback to facilitate mind-mindedness. For example, if a post did not mention the infant’s internal state (e.g., “We’re having breakfast”), the feedback would prompt the caregiver to consider the event in light of the infant’s internal states (e.g., “What does [name] like most for breakfast?”), reply in the first person from the infant’s perspective (e.g., “That’s yummy!”), or model the use of mental states (e.g., “Looks like they really like that yogurt!”).

The app was administered to a community sample of mothers who were recruited in the last trimester of pregnancy and began using the app as soon as the infant was born. These dyads were followed up when the infants were 6 months of age, at which point mothers’ mind-mindedness was assessed from a laboratory free-play mother–infant interaction. Mind-mindedness in the intervention group was compared with that in a control sample of mothers who were recruited when their infants were 6 months and had not received the smartphone app. Compared with the control group, mothers in the intervention group produced significantly more appropriate mind-related comments and significantly fewer non-attuned comments when interacting with their infants. Larkin et al. (2019) also investigated whether the app was equally effective in facilitating mind-mindedness in younger versus older mothers, given previous research findings indicating that teenage and younger mothers are less mind-minded than their older counterparts (Demers, Bernier, Tarabulsky, & Provost, 2010). The younger mothers who had received the app were more mind-minded than both younger and older mothers in the control group, suggesting that the app mitigated the impact of early motherhood on mind-mindedness. These findings are important because they provide evidence that mind-mindedness can be facilitated remotely using a technique that could be administered to parents on a grand scale.

2.2 Factors that may Moderate the Efficacy of Mind-Mindedness Interventions

Aside from maternal age (Larkin et al., 2019), previous research has not yet explored factors that may moderate the efficacy of an intervention that aims to facilitate mind-mindedness. We therefore do not know whether some mothers may be more receptive than others to procedures that encourage them to engage with and reflect on their infants’ internal states. Mind-mindedness is one of a number of constructs that fall under the umbrella term *parental mentalization* (see Zeegers, Colonnese, Stams, & Meins, 2017). Another aspect of parental mentalization is parental reflective functioning (PRF; Arietta Slade, 2005). While mind-mindedness focuses specifically on the caregiver’s mentalistic interpretations of the infant’s behavior during caregiver–infant interaction, PRF assesses the caregiver’s more general ability to mentalize about the caregiver–child relationship and themselves in the caregiving role.

PRF is typically measured from analyzing caregivers’ responses to interviews, such as the Parent Development Interview (PDI; Aber, Slade, Berger, Bresgi, & Kaplan, 1985), and is scored on an 11-point scale, ranging from bizarre/inappropriate (-1) to high/exceptional (+9). Higher scores indicate the caregiver’s ability to talk about the complex nature of internal states, the interplay between the caregiver’s internal states and those of their child, and the fact that internal states are variable and often difficult to read. Lower PRF has been reported to relate to a range of demographic and psychosocial risk factors (Arkle et al., 2023). Mothers who demonstrate more sophisticated engagement with and reflection on their infants’ thoughts and feelings may be likely to be receptive to an intervention that encourages them on a daily basis to think about what is on their infant’s mind. Levels of PRF at baseline may therefore moderate the effectiveness of the app in facilitating mind-mindedness.

The present study also included a number of other important variables that may relate to the efficacy of

the mind-mindedness intervention. Although the existing evidence suggests that mothers’ mental health is unrelated to mind-mindedness unless the mother is diagnosed with severe mental illness (Crucianelli et al., 2019; Meins, Fernyhough, Arnott, Turner, & Leekam, 2011; Schacht et al., 2017), higher levels of depression or anxiety or low levels of self-esteem or social support may impair the mother’s ability to engage with the app and benefit from the intervention. Similarly, while previous research indicates that infant temperament is unrelated to mind-mindedness (Larkin et al., 2019; Meins et al., 2011), having a temperamentally difficult infant may make it more difficult for the mother to benefit from the intervention. The present study therefore assessed mothers’ psychological wellbeing and infant temperament, and controlled for these variables in evaluating the effectiveness of the mind-mindedness app intervention.

2.3 The Present Study

The main aim of the present study was to evaluate the efficacy of the app used in Larkin et al.’s (2019) study using a rigorous randomized controlled trial design. We chose to assess mind-mindedness at two time points: baseline (when infants were age 6 months) and follow-up at age 12 months (the age at which the intervention or control procedures terminated). We predicted that, compared with the control group, the intervention group would achieve higher scores for appropriate mind-related comments and lower scores for non-attuned mind-related comments post-treatment. We also hypothesized that there would be a significant increase in appropriate mind-related comments from baseline to follow-up specifically in the intervention group. It was somewhat less straightforward to formulate a prediction regarding change over time for non-attuned comments. Individual differences in mind-mindedness – and perhaps particularly in non-attuned comments – arise in part because of the opacity of young infants’ internal states. For much of the first year, infants are not mobile and cannot talk, and therefore have little agency in demonstrating or expressing their desires, likes, dislikes, thought processes, and so on. But as infants become able to crawl, walk, and talk, it is arguably easier for caregivers to interpret their thoughts and feelings (Fishburn et al., 2022); non-attuned comments may thus naturally decline over the first year of life. We therefore predicted that both the intervention and control groups would show a decrease in non-attuned comments from baseline to follow-up, but that the decrease would be more marked in the intervention group. These hypotheses were made *a priori*, a power analysis was used to determine sample size, and the primary outcomes were pre-registered; we therefore consider our approach to be confirmatory. The study also explored whether mothers’ baseline PRF moderated the efficacy of the app intervention; these analyses were exploratory.

3 Method

3.1 Participants

Participants were 152 mothers with infants (72 girls) aged around 6 months ($M = 28.55$ weeks, $SD = 4.29$, range 21–48), who were recruited between June 2014 and June 2018. Participants lived in the cities of York and Hull and the surrounding areas in the north east of England. Recruitment was partly through local maternity services, whereby clinicians identified suitable candidates and referred them to the study team. The study team also recruited directly within the community (children’s centers, community events, parent–baby groups, social media, and word of mouth). Reflecting the ethnic makeup of the local area, 146 of the mothers were White; 135 were in a relationship with the infant’s father, 14 were single, and 30 were aged 22 or younger. Eligibility criteria were speaking to the infant in English, being aged 15 or older, and owning or having regular access to a smartphone. While the trial was initially designed to recruit only teenage and young mothers up to age 22, the upper age limit was removed 10 months after the study commenced to ensure that adequate numbers were reached to test the efficacy of the intervention. The study was approved by the National Health Service Health Research Authority Research Ethics Committee (REC reference: 14/NE/0114, IRAS project ID: 126036), and testing was conducted in line with the guidance of the British Psychological Society and American Psychological Association.

All participants provided written informed consent to participate. In line with National Health Service guidelines, parental consent was additionally required for participating mothers who were age 15. Participants were paid £20 and travel costs per testing session and received an additional £10 gift voucher if they completed the follow-up session at age 12 months.

3.2 Materials and methods

Overview of testing procedures. Testing took place either in the university laboratory (87%), at a community children’s center or school (11%), or in the participant’s home (1%). At age 6 months, participants completed the mind-mindedness assessment, an infant temperament measure, the PDI (to assess PRF), and questionnaires assessing depression and anxiety, self-esteem, and perceived social support. Mothers also provided demographic information. They indicated which level of education they had completed on a 0–6 scale: primary, secondary with/without qualifications, college with/without qualifications, occupational training, third level course, undergraduate degree, postgraduate degree. They detailed their living arrangements, sources of income, occupation, family composition, and provided information about their mental and physical health. After the testing procedures had been completed, mothers were randomly assigned to the intervention or control group (see below). Mothers and infants returned for the follow-up phase at age 12 months, and completed the mind-mindedness assessment for a second time. At this time point, mothers in both the intervention and control groups provided feedback on how often they had used the designated app according to the following scale: 0 = rarely or never, 1 = a few times a month, 2 = a few times a week, 3 = every day.

Mind-mindedness assessment. The procedure was identical at the 6 and 12 month phases. Mother–infant dyads were observed in a 10-minute free-play session. Mothers were instructed to play with the infant as they would do if they had free time together at home. The sessions were later transcribed verbatim and coded for mind-mindedness using Meins and Fernyhough’s (2015) criteria, in conjunction with viewing the filmed interaction. Comments in which the mother referred to the infant’s internal state or commented in the first person on the infant’s behalf (mind-related comments) were identified and each was coded dichotomously as appropriate or non-attuned. The former denotes comments where the attributed internal state (a) is consistent with the infant’s behavior at that moment in time (e.g., saying the infant wants a toy if he or she reaches toward it), (b) links the infant’s current internal state with related events in the past or future (e.g., “You’ll want to pet the lambs when we visit the farm tomorrow” while the infant played with a toy lamb), or (c) sought to engage the infant in play after a lull in the interaction (e.g., “You’ll like this rattle”). Comments where mothers voiced in the first person what the infants would be likely to say if they could talk were also coded as appropriate. Non-attuned mind-related comments were where the attributed internal state (a) is inappropriate given the infant’s ongoing behavior or seeks to redirect the infant to a new activity while the infant is actively engaged with a toy (e.g., “You don’t want to play with that anymore”), (b) is unrelated to the infant’s current activity (e.g., “Do you want to go out for lunch?” while the infant is playing with the stacking rings), (c) appears to project the mother’s internal state onto the infant, or (d) has an ambiguous referent.

All of the observations were coded for mind-mindedness by a trained researcher who was blind to all other data, including group allocation, with a randomly-selected 25% of sessions at each phase coded by a second trained, blind researcher. Coding of the follow-up sessions was conducted four years after coding the baseline sessions. Coders were not blind to the main hypothesis of the study (testing the efficacy of the intervention), but were blind to the hypothesized moderating effect of PRF. Inter-rater reliability for coding mind-related comments as appropriate or non-attuned was $\kappa = .74$; disagreements were resolved by discussion. To control for variation in maternal verbosity, scores for appropriate mind-related comments and non-attuned mind-related comments were expressed as a percentage of the total number of comments made during the session; percentage scores were used in all analyses.

PRF assessment. PRF was assessed from mothers’ responses to the PDI-R2-S (A. Slade, Aber, Bresgi, Berger, & Kaplan, 2004). This semi-structured interview explores caregivers’ representations of themselves as a parent, of their child, and of the parent–child relationship. Permission was given by the lead author (Slade, pers comm) to reduce the interview to 20 items by leaving out sections D (Parent’s Family History) and E (Separation and Loss). The interviews were audio recorded and transcribed verbatim. Using the Reflective Functioning Manual (Fonagy, Target, Steele, & Steele, 1998) and the Addendum to Reflective Functioning Scoring Manual (A. Slade, Bernbach, Grienberger, Levy, & Locker, 2004), a trained coder read the transcripts and assigned each an overall score for PRF. PRF is rated on an 11-point continuous scale from -1 (“Negative or Bizarre PRF”) to 9 (“Marked PRF”). A randomly selected subset of the transcripts ($N = 21$) was coded by a second trained, blind coder in order to measure inter-rater reliability; ICC = .76. Both coders were blind to the hypotheses of the study.

Infant temperament. At age 6 months, infants took part in the car seat task from the Infant Laboratory

Temperament Assessment Battery (Lab-TAB; Goldsmith & Rothbart, 1996) to assess their temperament. This task has strong ecological validity, as well as good predictive and concurrent validity in measuring temperament when used in isolation from the full LabTAB battery (Hay et al., 2010, 2014). The mother was asked to strap their infant into a car seat (fixed to an upright chair in front of a camera) without speaking to the infant. The mother was instructed then to stand to the side, slightly behind the car seat (so that the infant could see them by turning their head), and refrain from looking at the infant. Thirty seconds were timed from when the mother closed the buckle of the car seat.

The filmed 30-second period was later divided into six 5-second epochs, with each coded for the presence and intensity of behaviors indicative of frustration and sadness. A trained researcher who was blind to all other data coded all of the observations, with a second trained, blind rater coding a randomly selected 25%. Both coders were blind to the hypotheses of the study. Inter-rater reliability scores for the ratings were: facial anger: ICC = .74; facial sadness: ICC = .77; distress vocalization: ICC = .85, and physical struggle: ICC = .81. Scores across the epochs were averaged to give mean scores for facial anger, facial sadness, distress vocalization, and physical struggle. Scores for these four scales showed good internal reliability, $\alpha = .71$. Composite scores were therefore calculated and used in the analyses. Higher scores indicated greater negative affect, distress, and struggle, and index more difficult temperament.

Questionnaire measures. Maternal depression and anxiety were assessed using the Hospital Anxiety and Depression Scale (HADS; Zigmond & Snaith, 1983). The HADS is a 14-item questionnaire scored on a 4-point Likert scale. It has two subscales: anxiety (HADS-A), and depressive symptoms (HADS-D), which can be summed to give a total HADS score ranging between 0 and 42. A higher score indicates more anxiety/depressive symptoms; internal reliability, $\alpha = .79$. Total HADS scores were used in the analyses.

Self-esteem was assessed using the Rosenberg Self-Esteem Scale (Rosenberg, 1965). This 10-item self-report questionnaire generates a global measure of self-esteem, with each item rated on a 4-point Likert scale, providing a total score between 0 and 30. Higher scores indicate higher levels of self-esteem; internal reliability, $\alpha = .89$.

Perceived social support was assessed using the Multi-dimensional Scale of Perceived Social Support (Zimet, Dahlem, Zimet, & Farley, 1988). The 12-item questionnaire focuses on social isolation, loneliness, and whether individuals feel they have someone who can provide psychological support if needed. Items are rated using a 7-point Likert scale, yielding possible scores between 12 and 84, with higher scores indexing higher perceived social support; internal reliability, $\alpha = .95$.

Randomization procedure. Figure 1 shows the CONSORT flow diagram for the enrollment, allocation, and follow-up phases of the study. Participants were randomly allocated to the intervention ($n=73$) or control group ($n=78$) using the software Minim, which was programmed to account for maternal age to ensure that both groups contained approximately the same number of younger (19 and younger) and older (over 19) mothers. Randomization occurred on a rolling basis across the recruitment period, with recruitment, randomization, and data collection occurring in parallel. To avoid bias in data collection, a different researcher ran the software and wrote down the designated group while the testing was being completed. The allocated group was given to the researcher conducting the testing in a sealed envelope, which was opened when all of the testing had been completed. The researcher then administered the intervention or control procedure accordingly.

Intervention procedure. Mothers assigned to the intervention group were shown a 10-minute animated film commissioned and developed by the study team to introduce them to the concept of mind-mindedness. The experimenter then modeled what the participant's infant might be thinking or feeling (e.g., "I really like this toy", "I've never been here before"), emphasizing the importance of 'tuning in' to the infant's internal states. The mothers were asked to imagine what their infant might be thinking or feeling in various hypothetical situations, and indicated what they could say to show they understood the infant's internal states. The experimenter then introduced the BabyMind© app (see below) and installed it on the mother's smartphone. The experimenter demonstrated the different tabs and ensured that the mother knew how to use the app. Participants consented to being sent a gentle text reminder if they had not uploaded a post for a week, and to receiving a call if they had still not posted after a further 3 days. The procedure lasted around 20 minutes.

The research team designed the BabyMind© app in conjunction with IC Mobile Lab and in consultation with parents and maternity professionals. The app contained four tabs: My Baby, Alerts, Info, and Journal. Users entered their infant's date of birth, name, and gender, and uploaded a profile photograph in the My

CONSORT 2010 Flow Diagram

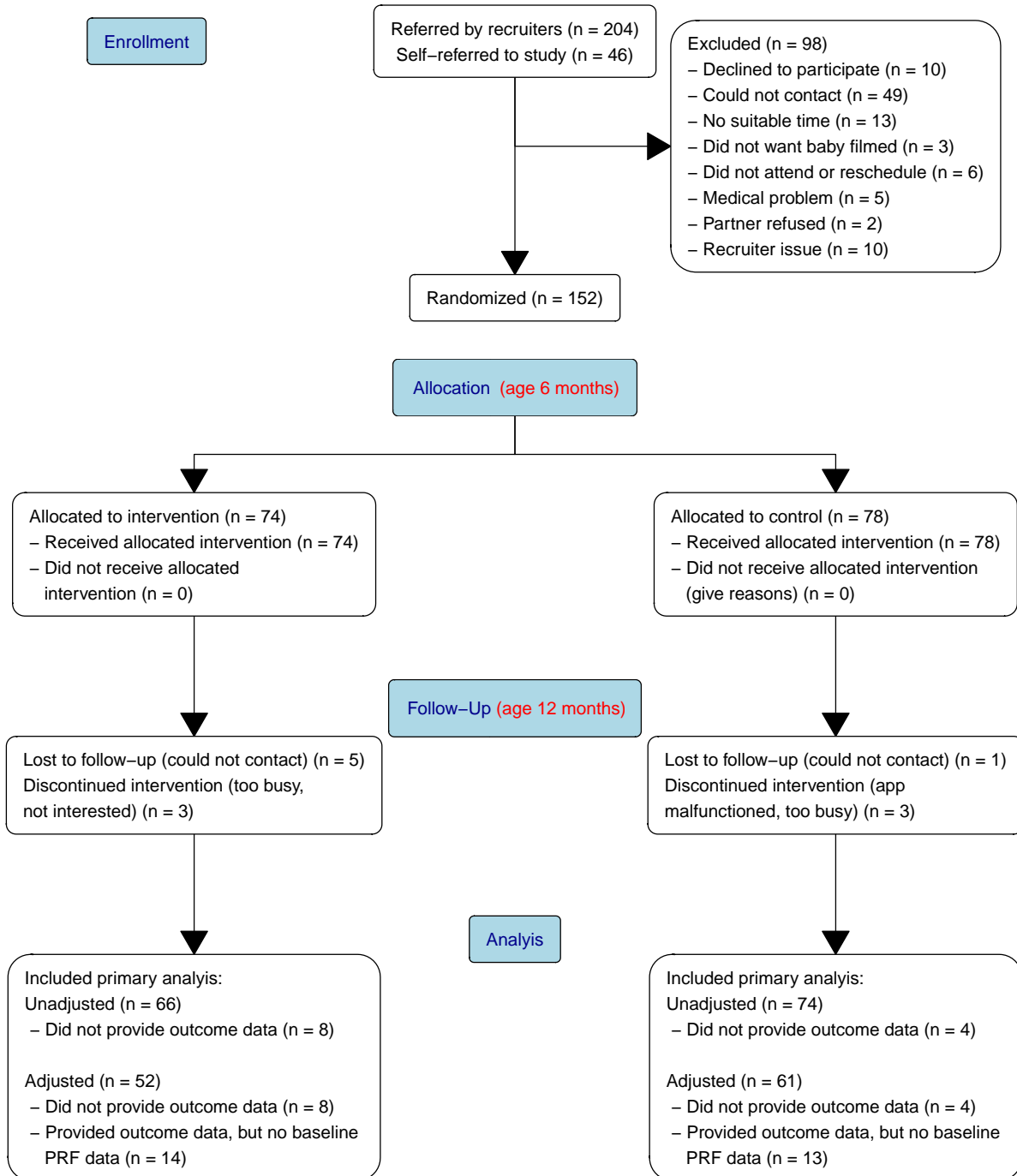


Figure 1: CONSORT flow diagram summarizing the design of the randomized controlled trial.

Baby tab.

Parents received daily alerts which provided brief research findings relating to infants’ psychological development via the Alerts tab. Alerts were programmed to contain the name of the user’s infant and were tailored to the age of the user’s infant in order to encourage mothers to appreciate their own infant’s cognitive, linguistic, or socio-emotional abilities; all alerts were stored within the tab for future reference. The user also received a weekly cartoon depicting an infant engaged in a common activity, with a caption that modeled what the infant might say if he or she could speak. The Alerts tab linked with the Info tab, which provided additional information relating to six areas of development: *Forming relationships*, *Learning about the world*, *Learning to talk*, *Tuning into a baby’s mind*, *Ideas for play*, and *How do we study babies?* The Alerts and Info tab were designed to support users to hold developmentally accurate expectations of their infants, and to provoke curiosity about their infants’ minds.

Users received the daily alert “What’s on [Name’s] mind?” via the Journal tab to prompt them to reflect on what the infant might be thinking or feeling. The user could upload a statement, photograph, or video clip in response to the alert. The tab’s text options included a wide range of emojis, and parents were informed that they could add these to their posts or simply post the emoji without any text. No guidance was given on the type of material that the users should upload. The research team viewed the uploaded material via a secure server, evaluated whether or not the post was mind-related, and responded accordingly. Users then viewed the research team’s response in the ‘Journal’ tab. This response was the end of the process; the app was designed to prevent extended interaction between the user and the research team in order to avoid introducing additional variance in the amount of feedback provided.

Control procedure. Mothers watched a 10-minute animated film on general infant development in the first year of life that did not cover mind-mindedness. The experimenter then asked the mother (a) whether they had noticed their baby doing any of the things from the film, (b) what changes they had noticed in their baby since birth, and (c) how they kept track of their baby’s development. The experimenter then introduced the control app. The app that was selected was WebMD©, which is freely available via the App Store. We chose this app because it was very similar to BabyMind© in all ways apart from the prompts specifically to reflect on the infant’s internal states. The control app had a Settings tab where the mother entered their infant’s profile information. The Home tab delivered daily tips and weekly features targeted to the age of the user’s infant. The Baby 101 tab provided information on infant health, growth, and developmental milestones, as well as tips on effective parenting. The Journal tab allowed the mother to chart their infant’s development by uploading photographs and comments, and mothers could also record how much their infants were feeding, sleeping, and growing. Mothers were instructed to use the app as much as possible over the following 6 months. The procedure lasted around 20 minutes.

3.3 Study Design and Data Analysis

The study was registered as a clinical trial with the ISRCTN Registry (<https://www.isrctn.com/ISRCTN19383681>), with the two mind-mindedness indices (appropriate and non-attuned mind-related comments) stipulated as the primary outcome measures. A statistical analysis plan was not pre-registered for this study. The study was conducted in alignment with the CONSORT statement for randomized trials of nonpharmacologic treatments (Boutron et al., 2017). As well as adopting this gold standard design, we used an active control condition (see above). A power calculation indicated that 67 participants per group were required to detect a medium-size effect with 0.8 power and $\alpha = 0.05$. The trial used a superiority design, based on the assumption that participants who received the intervention would be superior with regard to mind-mindedness compared with those who received the active control.

Statistical analysis was carried out using Stata version 17.0. Estimates of treatment effect are presented alongside 95% confidence intervals (CIs) and p values. Statistical significance testing was carried out at the 5% level.

Table 1: Descriptive Statistics for All Variables as a Function of Group

	BabyMind©	Control
	Mean (SD) Range	Mean (SD) Range
Maternal age in years	30.3 (5.99) 16–42	30.88 (7) 17–47
Infant age in weeks	28.99 (3.9) 22–40	28.15 (4.61) 21–48
Maternal education	5.42 (1.57) 2–7	5.35 (1.62) 2–7
HADS	8.7 (4.7) 0–19	9.47 (4.4) 0–22
Self-esteem	21.49 (4.3) 12–30	21.64 (4.86) 9–30
Perceived social support	71.42 (13.46) 12–84	71.68 (13.73) 12–84
Infant temperament	2.54 (2) 0–7.5	2.91 (2.31) 0–9.17
PRF	4.38 (1.34) 2–7	4.29 (1.37) 0–7
AMRC at 6 months (%)	3.7 (2.42) 0–10.85	4.24 (2.69) 0–11.59
NAMRC at 6 months (%)	2.37 (2.2) 0–8.46	2.53 (1.92) 0–8.93

Note: HADS = Hospital Anxiety and Depression Scale; PRF = Parental Reflective Functioning; AMRC = appropriate mind-related comments; NAMRC = non-attuned mind-related comments.

Table 2: Self-Reported App Usage as a Function of Group

	BabyMind©	Control
	Number (%)	Number (%)
Rarely or never	9 (13.6)	13 (17.6)
A few times a month	21 (31.8)	33 (44.6)
A few times a week	24 (36.4)	26 (35.1)
Every day	12 (18.2)	2 (2.7)

4 Results

4.1 Descriptive Statistics and Preliminary Analyses

PRF data were available for 123 participants. Administration instructions for the PDI-R2-S state that the infant must not be present during the interview; missing data were due to the interview being terminated because of infants needing attention. Temperament data were available for 142 infants; missing data were due to termination of the assessment because of high infant distress, the requisite testing equipment not being available, or the mother not following the protocol instructions. At follow-up (age 12 months), data were available for 140 mothers on the primary outcomes (appropriate and non-attuned mind-related comments), with 8 intervention group and 4 control group dyads lost to the study between baseline and follow up (see Figure 1). No harms were reported by intervention or control group. Feedback was sought on the interventions at follow up, and while occasional technical difficulties were reported, no distress or adverse events were reported.

Table 1 shows the descriptive statistics for all variables as a function of whether participants received the BabyMind© or control app. In line with advice on analyses relating to randomized controlled trials (e.g., Boer, Waterlander, Kuijper, Steenhuis, & Twisk, 2015), formal statistical testing of differences between the groups at baseline was not carried out.

Table 2 presents the usage data for the intervention and control groups. Adherence to the protocol was defined as using the app a few times a week or every day. As shown in Table 2, more mothers in the intervention group adhered to the protocol (36 out of 66) compared with their control group counterparts (28

Table 3: Bivariate Correlations between Variables at Baseline

	1	2	3	4	5	6	7	8	9
1. Maternal age									
2. Infant age	-0.21*								
3. Maternal education	0.67+	-0.22**							
4. HADS	-0.05	0.11	-0.07						
5. Self-esteem	0.1	-0.12	-0.01	-0.38+					
6. Perceived social support	0.28+	-0.15	0.29+	-0.3+	0.2*				
7. Infant temperament	0.04	-0.04	0.09	-0.05	0.11	0.16			
8. PRF	0.45+	-0.09	0.48+	0.09	-0.06	0.23*	0.08		
9. AMRC	0.14	0.08	0.15	-0.01	0.14	0.09	0.12	0.16	
10. NAMRC	-0.05	-0.04	-0.01	-0.16*	0.06	-0.1	-0.07	-0.14	0.09

* $p < .05$, ** $p < .01$, + $p < .001$.

Note: HADS = Hospital Anxiety and Depression Scale; PRF = Parental Reflective Functioning; AMRC = appropriate mind-related comments; NAMRC = non-attuned mind-related comments.

out of 74), $\chi^2(1) = 3.92$, $p = .048$.

Table 3 shows the correlation matrix for all study variables at baseline. As shown in Table 3, scores for the two mind-mindedness indices were unrelated to the demographic and maternal mental health variables, apart from the small-effect negative correlation between non-attuned mind-related comments and HADS scores. These findings replicate those in the extant literature (e.g., Meins et al., 2011). Replicating previous null findings (e.g., Meins et al., 2012), scores for appropriate and non-attuned mind-related comments were unrelated at baseline (see Table 3); the two mind-mindedness indices were also unrelated at age 12 months, $r(138) = -0.13$, $p = .117$. At baseline, PRF scores were unrelated to both indices of mind-mindedness (see Table 3), replicating previous null findings (Dollberg, 2022).

4.2 Evaluating the Efficacy of the BabyMind© App

Scores for appropriate and non-attuned mind-related comments at age 12 months are shown in Table 4. The primary outcomes of percentage of appropriate mind-related comments and percentage of non-attuned mind-related comments were each analyzed using two different methods. The unadjusted mean difference between treatment groups was estimated using a t test. Note that this analysis used only post-intervention observed data, with the 12 participants who did not provide outcome data deleted, and did not include covariates. The unadjusted mean difference between the intervention and control groups was significant for appropriate mind-related comments, $t(138) = 2.84$, $p = .005$, 95% CI 0.48, 2.67, $d = .48$, but there was no significant group difference for non-attuned mind-related comments, $t(138) = 1.96$, $p = .052$, 95% CI 0, -1.08, $d = -.33$. Next, the adjusted mean difference was estimated using linear regression, adjusting for PRF, baseline percentage of appropriate mind-related comments, and baseline percentage of non-attuned mind-related comments as fixed effects. This analysis also used only observed data; there were an additional 27 participants (Intervention 14; Control 13) with primary outcome data who were excluded from the adjusted analyses due to not providing data on PRF at baseline. Table 4 summarizes the results of these regression analyses. As shown in Table 4, the adjusted mean group differences for both appropriate mind-related comments and non-attuned mind-related comments were significant. Compared with their control group counterparts, intervention group mothers scored significantly higher for appropriate mind-related comments and significantly lower for non-attuned mind-related comments at age 12 months. The adjusted Cohen's d effect sizes for appropriate mind-related comments and non-attuned mind-related comments were .61 and -.55 respectively.

The change over time in the two mind-mindedness indices for the intervention and control groups is shown in Figures 2 and 3 and was tested using paired samples t tests. In the intervention group, there was a significant increase in appropriate mind-related comments, $t(65) = 2.41$, $p = .019$, $d = 0.3$, 95% CI: 0.06, 0.54, and a significant decrease in non-attuned mind-related comments, $t(65) = 5.59$, $p < .001$, $d = -0.69$, 95%

Table 4: Results of the Analyses Testing the Efficacy of the BabyMind© App on the Primary Outcomes at Age 12 Months Follow-up

	AMRC 12 months				NAMRC 12 months			
	B	SE	95% CI	p value	B	SE	95% CI	p value
Group	1.99	0.54	0.92, 3.06	<.001	-0.89	0.3	-1.48, -0.3	.004
PRF	-0.02	0.21	-0.43, 0.39	.91	0.03	0.12	-0.19, 0.26	.766
Baseline AMRC	0.47	0.11	0.26, 0.69	<.001	-0.05	0.06	-0.17, 0.06	.374
Baseline NAMRC	0.31	0.14	0.03, 0.59	.029	0.24	0.08	0.09, 0.4	.003
	F(4, 108)	= 9.09,	p < .001,	R ² = .25	F(4, 108)	= 4.37,	p = .003,	R ² = .14
Note: Included in the model n = 113; Reference category = Control; AMRC = appropriate mind-related comments (%); NAMRC = non-attuned mind-related comments (%); PRF = Parental Reflective Functioning.								

CI: -0.48, -0.94. In the control group, both indices showed significant decreases over time: for appropriate mind-related comments $t(73) = 3.12$, $p = .003$, $d = -0.36$, 95% CI: -0.14, -0.6; for non-attuned comments, $t(73) = 3.2$, $p = .002$, $d = -0.37$, 95% CI: -0.15, -0.61.

4.3 Sensitivity Analyses

Missing data. The sensitivity of the primary outcome analyses to missing data was assessed. Baseline characteristics of participants who did and did not provide follow-up data were compared using non-parametric tests (Mann-Whitney U Test for continuous variables and Fisher’s Exact Test for categorical variables). Statistically significant differences were found for PRF ($p = .041$), infant temperament ($p = .033$), and maternal education ($p = .019$), with mothers who provided follow-up data scoring more highly for PRF and educational level, and having infants with more difficult temperaments.

Missing values of baseline covariates were replaced with the mean value; missing data for the other variables was imputed using multiple imputation by chained equations (White, Royston, & Wood, 2011). The imputation model for each primary outcome included the variables used in the adjusted primary analyses (PRF, baseline percentage of appropriate comments, baseline percentage of non-attuned comments, and treatment allocation), along with baseline variables thought a priori to be predictive of missingness (maternal age, infant age, highest level of maternal education, number of children, HADS, self-esteem, perceived social support, and infant temperament).

One hundred imputed datasets were generated in order to aid reproducibility of the missing data analyses. Each imputed dataset was analyzed using a t test, and the estimates obtained were combined using Rubin’s rules (White et al., 2011). Each imputed dataset was also analyzed using linear regression, adjusting for PRF, baseline percentage of appropriate comments, and baseline percentage of non-attuned comments as fixed effects, with estimates also combined using Rubin’s rules.

The sensitivity analyses did not affect the interpretation of the results. The unadjusted mean difference obtained using multiple imputation by chained equations for appropriate mind-related comments was 0.34, 95% CI: -5.64 to 6.31, $p = 0.912$, and the adjusted mean difference was 1.95, 95% CI: 0.84 to 3.05, $p = 0.001$. For non-attuned mind-related comments, the unadjusted mean difference was -0.88, 95% CI: -3.49 to 1.73, $p = 0.507$, and the adjusted mean difference was -0.83, 95% CI: -1.45 to -0.21, $p = 0.009$.

Adherence. A complier-average-causal-effect (CACE) analysis was carried out (Hernan & Robins, 2020; Hewitt, Torgerson, & Miles, 2006) in order to estimate the treatment effect for those participants who would adhere to the BabyMind© app intervention if offered it. The CACE analysis was carried out using a two-stage least squares instrumental variable regression model. The instrumental variable was treatment allocation, with adherence being the endogenous variable, and PRF, baseline percentage of appropriate mind-related comments, and baseline percentage of non-attuned mind-related comments being the exogenous variables. A participant was classed as adhering to the intervention if they reported using the app either a few times a week or every day. The results of the CACE analysis estimated the treatment effect in adherers to be 2.61, 95% CI: 1.3 to 3.91, $p < .001$, for appropriate mind-related comments, and -1.02, 95% CI: -1.8 to -0.24, $p = .011$, for non-attuned mind-related comments.

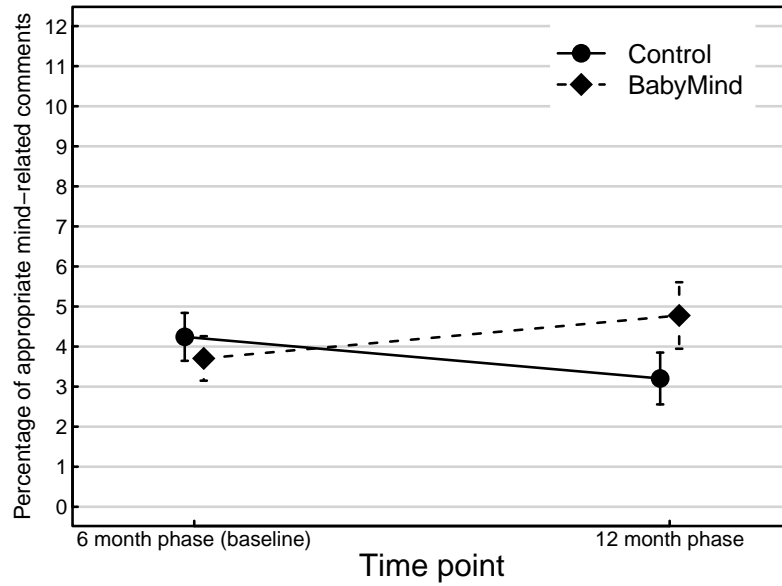


Figure 2: Mean value of the percentage of appropriate mind-related comments at the 6-month (baseline) and 12-month phases as a function of group. The error bars represent 95% confidence intervals.

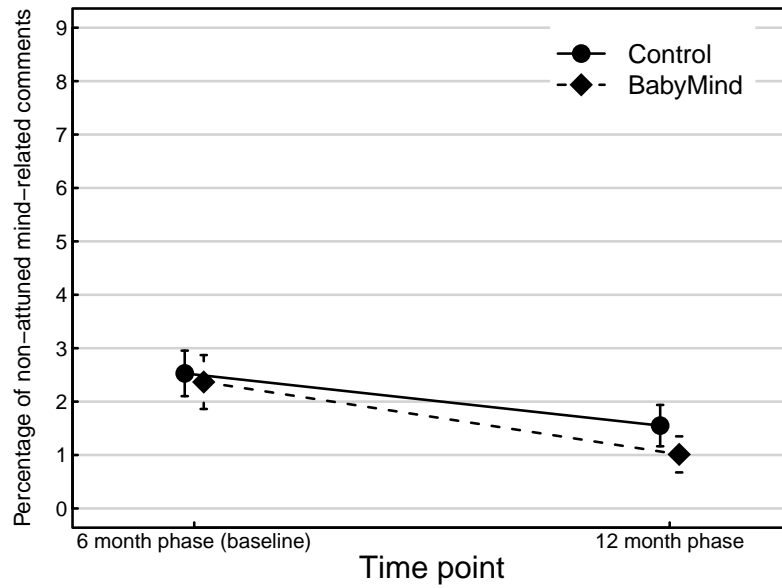


Figure 3: Mean value of the percentage of non-attuned mind-related comments at the 6-month (baseline) and 12-month phases as a function of group. The error bars represent 95% confidence intervals.

4.4 Does PRF Moderate the Efficacy of the Intervention?

To investigate whether PRF moderated the efficacy of the intervention, an exploratory analysis was carried out for each mind-mindedness index via the addition of an interaction term between PRF and treatment group to the linear regression equation described above (Brankovic et al., 2019). This model was then compared to the primary analysis model using a likelihood ratio test. There was no evidence of an interaction between treatment group and PRF for either appropriate mind-related comments ($p = .971$) or non-attuned mind-related comments ($p = .915$).

5 Discussion

The main aim of the study reported here was to investigate whether the BabyMind© smartphone app was effective in facilitating mind-mindedness when evaluated using a rigorous randomized controlled trial design. Adjusting for baseline PRF, appropriate mind-related comments, and non-attuned mind-related comments, mothers who had received the app scored more highly than their control group counterparts for appropriate mind-related comments at follow-up (age 12 months). The effect sizes for the impact of the intervention on both indices of mind-mindedness were medium to large (d s of .61 and -.55 for appropriate and non-attuned comments respectively). Given that mind-mindedness is characterized in terms of tending to comment appropriately on the infant's internal states while avoiding comments that misinterpret the infant's thoughts and feelings, these findings suggest that mind-mindedness can be successfully facilitated using the app. With regard to change over time, the results supported the hypothesized increase in appropriate mind-related comments specifically in the mothers who had received the BabyMind© app. In contrast, there was a decrease in appropriate comments from 6 to 12 months in mothers who had received the control app. In line with our proposal that non-attuned comments should decline over time because infants' internal states become more transparent by the end of the first year of life, both groups showed a decrease in non-attuned comments. However, the decrease was more marked in the intervention group, suggesting that the app effected an additional decrease in non-attuned comments over and above the expected natural decline by age 12 months.

To evaluate whether missing data had influenced the results of the trial, sensitivity analyses were conducted using multiple imputation, with the imputation model additionally including the demographic and mental health control variables, and infant temperament. The sensitivity analyses indicated that the BabyMind© app was more effective than the active control in relation to both appropriate and non-attuned comments when missing data were taken into account. We also conducted a formal analysis to establish the treatment effect of the BabyMind© app specifically in participants who would adhere to the intervention if offered it. This analysis indicated a significant treatment effect in adherers resulting in higher levels of appropriate mind-related comments and lower levels of non-attuned comments.

Finally, we explored whether mothers' PRF at baseline moderated the efficacy of the BabyMind© app in facilitating mind-mindedness. The analyses indicated that PRF did not moderate the app's efficacy in relation to either appropriate or non-attuned mind-related comments; including the interaction term between PRF and treatment group did not result in significant improvements to the models. These findings indicate that caregivers' tendency to show high levels of reflective functioning when discussing their child, themselves as parents, and the parent-child relationship during an interview does not make them more receptive to an intervention that aims to facilitate their engagement with and reflection on their infants' thoughts and feelings.

Bakermans-Kranenburg, van Ijzendoorn, and Juffer's (2003) meta-analysis provides useful data for gauging the effectiveness of our intervention in the broader context of parenting interventions. This meta-analysis investigated the effect of interventions that focused (a) specifically on facilitating caregiver sensitivity, (b) provided caregivers with social support and practical advice, (c) facilitated caregivers' construction of more optimal representations of themselves in the caregiver-child relationship, or (d) involved a combination of these approaches (e.g., sensitivity and support). Interventions that focused on sensitivity only were reported to be more effective in facilitating sensitive caregiving ($d = .45$) than all of the other types of intervention ($d = .27$). The meta-analytic data also showed that interventions starting when infants were older ($d = .44$) were more effective than those beginning in the first 6 months of life ($d = .28$). This study also tested the effectiveness of a small number of interventions that did not involve personal contact; the effect size for

the four studies in this category was $d = .62$. Together with the findings of the present study, these results suggest that effective intervention can be delivered without incurring the substantial costs associated with face-to-face procedures.

It is also informative to compare the results of the present study with those of Larkin et al. (2019), who evaluated the efficacy of the BabyMind© app in a separate sample of mothers who began using the app from birth and were followed up when their infants were age 6 months. In this previous study, the average mind-mindedness scores for mothers who had received the app intervention were 8.94% for appropriate mind-related comments and 0.77% for non-attuned comments. These figures can be compared with those of the intervention group in the present study at follow up: 4.78% for appropriate comments and 1.01% for non-attuned comments. Thus, while the present randomized controlled trial demonstrated that the app was effective in facilitating mind-mindedness, the mothers who had received the intervention in the present study appear to be less mind-minded, particularly in terms of appropriate mind-related comments, than those in Larkin et al.'s previous study.

One obvious difference between the two studies is the age at which mind-mindedness was assessed, which may account for the observed differences. Recall that the control group in the present study showed a decline in appropriate mind-related comments from 6 to 12 months. The BabyMind© app appeared to reverse this decline. A previous study also reported a decrease in mothers' appropriate mind-related comments over a somewhat longer period, from 7 to 19 months (McMahon, Camberis, Berry, & Gibson, 2016). The greater agency shown by older infants in terms of their linguistic and motor abilities may thus mean that mothers increasingly tend not to comment in either appropriate or non-attuned ways on what their infants are thinking or feeling because of the greater transparency of infants' internal states. The observational indices of appropriate and non-attuned comments have therefore been proposed to be less valid measures of mind-mindedness once infants have acquired these abilities (Fishburn et al., 2022). Indeed, the mind-mindedness coding manual (Meins & Fernyhough, 2015) recommends using the observational procedure specifically in the first year of life.

An alternative possibility for the difference across the two studies is that the app is more effective the earlier in the infant's life it is administered, in contrast to Bakermans-Kranenburg et al.'s (2003) findings on other parenting interventions. Young infants' states are difficult to read, and by providing new mothers with psychoeducation and prompts to consider the world from the infant's point of view at a time when they may be struggling to know how to approach caregiving, the app may be particularly effective in scaffolding mind-minded interaction. Future research should thus investigate whether early intervention with the BabyMind© app is more effective than later intervention. This research could also explore whether any such benefits of early intervention are more likely to be maintained over the course of infancy.

While our results provide robust evidence for the efficacy of the BabyMind© app in facilitating mind-mindedness, it is important to recognize the present study's limitations. Our participants varied in terms of SES and maternal age, but the vast majority were in a relationship with the baby's father, and our sample was overwhelmingly White, reflecting the ethnicity of the locality from which they were recruited. There is long-standing research highlighting how parenting practices and views on parenting vary in parents from different cultural and ethnic backgrounds (e.g., Choi, Kim, Kim, & Park, 2013), and research is beginning to address cultural and ethnic differences in parental mentalization. For example, Dai, McMahon, & Lim (2019) investigated mind-mindedness in Australian and Chinese mothers, and reported that Australian mothers made more appropriate mind-related comments and fewer non-attuned comments than their Chinese counterparts. Research on South Korean mothers has shown that they report high levels of certainty about their infants' internal states (Lee, Meins, & Larkin, 2021). Investigating whether the BabyMind© app is equally effective in facilitating mind-mindedness in different cultural and ethnic groups is therefore a fruitful avenue for future research. It would also be interesting to investigate whether the app effectively facilitates mind-mindedness in fathers, single parents, and other caregivers. A further limitation of the present study is the fact that mothers in the intervention group were more likely to adhere to the protocol than were their control group counterparts. While this suggests that mothers found the BabyMind© app engaging, it does raise the possibility that the intervention was effective partly because mothers in the intervention group adhered more to the procedure. That said, the CACE analysis estimated the treatment effect for those participants who would adhere to the BabyMind© app intervention if offered it.

As well as investigating the efficacy of the BabyMind© app in more diverse populations, future research could investigate whether the app is effective in facilitating mind-mindedness across more varied types of

interaction. Mind-mindedness is typically assessed from free-play caregiver–infant interactions rather than in more challenging contexts that may involve conflict between the caregiver and child or infant distress. An exception is a study by Miller, Kim, Boldt, Goffin, & Kochanska (2019) who measured mind-mindedness during a snack observation, in which the parent had to feed the infant and wash their face and hands. It would be interesting to explore whether parents who have received the app intervention are more mind-minded during everyday caregiving activities such as feeding, changing, getting the child ready to go out, and so on. Knowing about the infant’s desires, interests, and preferences in these contexts may be critical in enabling the caregiver successfully to negotiate difficult situations.

Previous research has demonstrated the efficacy of an individualized video-feedback intervention in facilitating mind-mindedness in mothers hospitalized due to severe mental illness (Schacht et al., 2017). Future research could explore whether the BabyMind© app is effective in facilitating mind-mindedness in mothers diagnosed with severe mental illness. It would also be useful to investigate the app’s efficacy in mothers in the community with less severe levels of perinatal mental illness. The fact that the app intervention involves regular interaction with mothers without the need for in-person appointments or home visits could make the app a useful clinical tool to supplement services offered to parents in the perinatal period.

Mind-mindedness is known to predict a wide range of positive aspects of children’s development (Aldrich et al., 2021; McMahon & Bernier, 2017). It is therefore important for future research to establish whether administration of the BabyMind© app can effect improvement in core developmental outcomes. It may be that differences between our intervention and control groups will be observed with respect to aspects of child development that are known to be predicted by early mind-mindedness (e.g., behavioral regulation, theory of mind). We are currently following up this sample of families to explore this question.

As well trialling the efficacy of the BabyMind© app, the present study also provides data on how the two parental mentalization constructs—mind-mindedness and PRF—relate to one another. When the two constructs were assessed concurrently at age 6 months, PRF was unrelated to both indices of mind-mindedness (appropriate and non-attuned mind-related comments), with small effect sizes for these correlations. PRF at baseline was similarly unrelated to both appropriate and non-attuned comments at follow-up (age 12 months), and these null findings were observed in mothers who had received the BabyMind© intervention and those in the control group. These results replicate the null associations between mind-mindedness and PRF reported recently by Dollberg (2022). This lack of association between mind-mindedness and PRF is interesting, since it suggests that, although both constructs fall under the umbrella term *parental mentalization*, they are empirically distinct aspects of caregivers’ representations of their infants’ minds. Mind-mindedness is a construct at the interface between caregivers’ representations of their infants’ internal states and infant–caregiver behavior. While the ability to engage with and reflect on the infant’s thoughts and feelings—an ability likely to be equivalent to PRF—is necessary for being mind-minded, our results indicate that this capacity is not sufficient for the caregiver to demonstrate mind-mindedness when interacting with their infant. These null results mirror the pattern of findings reported for the relation between parents’ theory of mind abilities and their mind-mindedness in relation to their child (Barreto, Fearon, Osório, Meins, & Martins, 2015; Devine & Hughes, 2019), and suggest that mind-mindedness cannot be reduced to caregivers’ underlying mentalizing abilities.

In summary, the results of the present study indicate that it is possible to facilitate mind-mindedness remotely via a smartphone app. Our use of a randomized controlled trial, involving an active control condition that was closely matched to the intervention, with sensitivity analyses accounting for missing data, provide assurance that our results are robust. The fact that BabyMind© delivers an effective intervention without the need for face-to-face sessions is noteworthy, particularly given problems in accessing services and support as a result of the COVID-19 pandemic. Future research should thus explore the feasibility of delivering BabyMind© as part of standard practice in maternity and early years services. Investigating whether BabyMind© is effective in facilitating mind-mindedness in more diverse groups of caregivers—and whether this intervention has a sustained positive impact on children’s development—are interesting avenues for future research.

6 Acknowledgements

This research was supported by grants ES/K010719/1 and ES/R004706/1 from the Economic and Social Research Council. We thank Karin Fothergill for her invaluable contribution to the study, and the families

for their generous participation. We acknowledge the role of many professionals in supporting recruitment: the research midwife teams in York, Harrogate, Leeds, and North Tyneside, Gleanna Oland, Olivia Bowes, Caroline Porter, Julie Stamper, and the schools and children’s centers involved. Thanks to Sally Johns and Sue Final for support with study approvals, and to Lee McLaughlin and Kush Mishra for app development and support. Data collection, transcription, and coding were supported by Corinne Lee, Georgina Seymour, and our undergraduate and postgraduate research volunteers. The data and analytic code necessary to reproduce the analyses presented here are publicly available, together with a computationally reproducible version, via the following link: <https://github.com/bakerdh/testbabymind>. The materials necessary to attempt to replicate the findings presented here can be obtained from the corresponding author. The main analyses to test the efficacy of the intervention were pre-registered: <https://www.isrctn.com/ISRCTN1938368>.

References

- Aber, J. L., Slade, A., Berger, B., Bresgi, I., & Kaplan, M. (1985). *The parent development interview: Interview protocol*. Unpublished manuscript: Barnard College, Columbia University, New York.
- Aldrich, N. J., Chen, J., & Alfieri, L. (2021). Evaluating associations between parental mind-mindedness and children’s developmental capacities through meta-analysis. *Developmental Review*, 60, 100946. <https://doi.org/10.1016/j.dr.2021.100946>
- Arkle, P., Larkin, F., Wang, Y., Lee, Y., Fernandez, A., Li, L. Y., & Meins, E. (2023). Early psychosocial risk factors and postnatal parental reflective functioning. *Infancy*, 28(5), 864–881. <https://doi.org/10.1111/inf.12552>
- Bakermans-Kranenburg, M. J., IJzendoorn, M. H. van, & Juffer, F. (2003). Less is more: Meta-analyses of sensitivity and attachment interventions in early childhood. *Psychol Bull*, 129(2), 195–215. <https://doi.org/10.1037/0033-2909.129.2.195>
- Barreto, A. L., Fearon, R. M. P., Osório, A., Meins, E., & Martins, C. (2015). Are adult mentalizing abilities associated with mind-mindedness? *International Journal of Behavioral Development*, 40(4), 296–301. <https://doi.org/10.1177/0165025415616200>
- Boer, M. R. de, Waterlander, W. E., Kuijper, L. D. J., Steenhuis, I. H. M., & Twisk, J. W. R. (2015). Testing for baseline differences in randomized controlled trials: An unhealthy research behavior that is hard to eradicate. *Int J Behav Nutr Phys Act*, 12, 4. <https://doi.org/10.1186/s12966-015-0162-z>
- Boutron, I., Altman, D. G., Moher, D., Schulz, K. F., Ravaud, P., & CONSORT NPT Group. (2017). CONSORT statement for randomized trials of nonpharmacologic treatments: A 2017 update and a CONSORT extension for nonpharmacologic trial abstracts. *Ann Intern Med*, 167(1), 40–47. <https://doi.org/10.7326/M17-0046>
- Brankovic, M., Kardys, I., Steyerberg, E. W., Lemeshow, S., Markovic, M., Rizopoulos, D., & Boersma, E. (2019). Understanding of interaction (subgroup) analysis in clinical trials. *Eur J Clin Invest*, 49(8), e13145. <https://doi.org/10.1111/eci.13145>
- Centifanti, L. C. M., Meins, E., & Fernyhough, C. (2016). Callous-unemotional traits and impulsivity: Distinct longitudinal relations with mind-mindedness and understanding of others. *J Child Psychol Psychiatry*, 57(1), 84–92. <https://doi.org/10.1111/jcpp.12445>
- Choi, Y., Kim, Y. S., Kim, S. Y., & Park, I. K. (2013). Is asian american parenting controlling and harsh? Empirical testing of relationships between korean american and western parenting measures. *Asian Am J Psychol*, 4(1), 19–29. <https://doi.org/10.1037/a0031220>
- Colonnesi, C., Wissink, I. B., Noom, M. J., Asscher, J. J., Hoeve, M., Stams, G. J. J. M., ... Kellaert-Knol, M. G. (2012). Basic trust: An attachment-oriented intervention based on mind-mindedness in adoptive families. *Research on Social Work Practice*, 23(2), 179–188. <https://doi.org/10.1177/1049731512469301>
- Crucianelli, L., Wheatley, L., Filippetti, M. L., Jenkinson, P. M., Kirk, E., & Fotopoulou, A. K. (2019). The mindedness of maternal touch: An investigation of maternal mind-mindedness and mother-infant touch interactions. *Dev Cogn Neurosci*, 35, 47–56. <https://doi.org/10.1016/j.dcn.2018.01.010>
- Dai, Q., McMahon, C., & Lim, A. K. (2019). Cross-cultural comparison of maternal mind-mindedness among australian and chinese mothers. *International Journal of Behavioral Development*, 44(4), 365–370. <https://doi.org/10.1177/0165025419874133>
- Demers, I., Bernier, A., Tarabulsky, G. M., & Provost, M. A. (2010). Mind-mindedness in adult and adolescent

- mothers: Relations to maternal sensitivity and infant attachment. *International Journal of Behavioral Development*, 34(6), 529–537. <https://doi.org/10.1177/0165025410365802>
- Devine, R. T., & Hughes, C. (2019). Let's talk: Parents' mental talk (not mind-mindedness or mindreading capacity) predicts children's false belief understanding. *Child Dev*, 90(4), 1236–1253. <https://doi.org/10.1111/cdev.12990>
- Dollberg, D. G. (2022). Mothers' parental mentalization, attachment dimensions and mother-infant relational patterns. *Attach Hum Dev*, 24(2), 189–207. <https://doi.org/10.1080/14616734.2021.1901297>
- Fishburn, S., Meins, E., Fernyhough, C., Centifanti, L. C. M., & Larkin, F. (2022). Explaining the relation between early mind-mindedness and children's mentalizing abilities: The development of an observational preschool assessment. *Dev Psychol*, 58(1), 17–31. <https://doi.org/10.1037/dev0001272>
- Fonagy, P., Target, M., Steele, H., & Steele, M. (1998). *Reflective-functioning manual, version 5.0, for application to adult attachment interviews*. Unpublished manuscript. London: University College London.
- Goldsmith, H. H., & Rothbart, M. K. (1996). *Prelocomotor and locomotor laboratory temperament assessment battery, lab-TAB; version 3.0*. Unpublished technical manual, Department of Psychology, University of Wisconsin, Madison, WI.
- Hay, D. F., Perra, O., Hudson, K., Waters, C. S., Mundy, L., Phillips, R., ... CCDS Team. (2010). Identifying early signs of aggression: Psychometric properties of the cardiff infant contentiousness scale. *Aggress Behav*, 36(6), 351–357. <https://doi.org/10.1002/ab.20363>
- Hay, D. F., Waters, C. S., Perra, O., Swift, N., Kairis, V., Phillips, R., ... Goozen, S. van. (2014). Precursors to aggression are evident by 6 months of age. *Dev Sci*, 17(3), 471–480. <https://doi.org/10.1111/desc.12133>
- Hernan, M. A., & Robins, J. M. (2020). *Causal inference: What if*. Boca Raton: Chapman & Hall/CRC.
- Hewitt, C. E., Torgerson, D. J., & Miles, J. N. V. (2006). Is there another way to take account of noncompliance in randomized controlled trials? *CMAJ*, 175(4), 347. <https://doi.org/10.1503/cmaj.051625>
- Kirk, E., Pine, K., Wheatley, L., Howlett, N., Schulz, J., & Fletcher, B. C. (2015). A longitudinal investigation of the relationship between maternal mind-mindedness and theory of mind. *Br J Dev Psychol*, 33(4), 434–445. <https://doi.org/10.1111/bjdp.12104>
- Laranjo, J., Bernier, A., Meins, E., & Carlson, S. M. (2010). Early manifestations of children's theory of mind: The roles of maternal mind-mindedness and infant security of attachment. *Infancy*, 15(3), 300–323. <https://doi.org/10.1111/j.1532-7078.2009.00014.x>
- Laranjo, J., Bernier, A., Meins, E., & Carlson, S. M. (2014). The roles of maternal mind-mindedness and infant security of attachment in predicting preschoolers' understanding of visual perspective taking and false belief. *J Exp Child Psychol*, 125, 48–62. <https://doi.org/10.1016/j.jecp.2014.02.005>
- Larkin, F., Oostenbroek, J., Lee, Y., Hayward, E., & Meins, E. (2019). Proof of concept of a smartphone app to support delivery of an intervention to facilitate mothers' mind-mindedness. *PLoS One*, 14(8), e0220948. <https://doi.org/10.1371/journal.pone.0220948>
- Lee, Y., Meins, E., & Larkin, F. (2021). Translation and preliminary validation of a korean version of the parental reflective functioning questionnaire. *Infant Ment Health J*, 42(1), 47–59. <https://doi.org/10.1002/imhj.21883>
- Lundy, B. L. (2003). Father- and mother-infant face-to-face interactions: Differences in mind-related comments and infant attachment? *Infant Behavior and Development*, 26(2), 200–212. [https://doi.org/10.1016/S0163-6383\(03\)00017-1](https://doi.org/10.1016/S0163-6383(03)00017-1)
- McMahon, C. A., & Bernier, A. (2017). Twenty years of research on parental mind-mindedness: Empirical findings, theoretical and methodological challenges, and new directions. *Developmental Review*, 46, 54–80. <https://doi.org/10.1016/j.dr.2017.07.001>
- McMahon, C. A., Camberis, A.-L., Berry, S., & Gibson, F. (2016). Maternal mind-mindedness: Relations with maternal-fetal attachment and stability in the first two years of life: Findings from an australian prospective study. *Infant Ment Health J*, 37(1), 17–28. <https://doi.org/10.1002/imhj.21548>
- Meins, E. (1997). *Security of attachment and the social development of cognition*. Hove: Lawrence Erlbaum Associates.
- Meins, E., Bureau, J.-F., & Fernyhough, C. (2018). Mother-child attachment from infancy to the preschool years: Predicting security and stability. *Child Dev*, 89(3), 1022–1038. <https://doi.org/10.1111/cdev.12778>
- Meins, E., Centifanti, L. C. M., Fernyhough, C., & Fishburn, S. (2013). Maternal mind-mindedness and children's behavioral difficulties: Mitigating the impact of low socioeconomic status. *J Abnorm Child Psychol*, 41(4), 543–553. <https://doi.org/10.1007/s10802-012-9699-3>

- Meins, E., & Fernyhough, C. (2015). *Mind-mindedness coding manual, version 2.2*. Unpublished manuscript. University of York, York, UK.
- Meins, E., Fernyhough, C., Arnott, B., Turner, M., & Leekam, S. R. (2011). Mother- versus infant-centered correlates of maternal mind-mindedness in the first year of life. *Infancy*, 16(2), 137–165. <https://doi.org/10.1111/j.1532-7078.2010.00039.x>
- Meins, E., Fernyhough, C., & Centifanti, L. C. M. (2019). Mothers' early mind-mindedness predicts educational attainment in socially and economically disadvantaged british children. *Child Dev*, 90(4), e454–e467. <https://doi.org/10.1111/cdev.13028>
- Meins, E., Fernyhough, C., Fradley, E., & Tuckey, M. (2001). Rethinking maternal sensitivity: Mothers' comments on infants' mental processes predict security of attachment at 12 months. *The Journal of Child Psychology and Psychiatry and Allied Disciplines*, 42(5), 637–648. <https://doi.org/10.1017/S0021963001007302>
- Meins, E., Fernyhough, C., Rosnay, M. de, Arnott, B., Leekam, S. R., & Turner, M. (2012). Mind-mindedness as a multidimensional construct: Appropriate and nonattuned mind-related comments independently predict infant-mother attachment in a socially diverse sample. *Infancy*, 17(4), 393–415. <https://doi.org/10.1111/j.1532-7078.2011.00087.x>
- Miller, J. E., Kim, S., Boldt, L. J., Goffin, K. C., & Kochanska, G. (2019). Long-term sequelae of mothers' and fathers' mind-mindedness in infancy: A developmental path to children's attachment at age 10. *Dev Psychol*, 55(4), 675–686. <https://doi.org/10.1037/dev0000660>
- Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton, NJ: University Press.
- Schacht, R., Meins, E., Fernyhough, C., Centifanti, L. C. M., Bureau, J.-F., & Pawlby, S. (2017). Proof of concept of a mind-mindedness intervention for mothers hospitalized for severe mental illness. *Dev Psychopathol*, 29(2), 555–564. <https://doi.org/10.1017/S0954579417000177>
- Slade, Arietta. (2005). Parental reflective functioning: An introduction. *Attach Hum Dev*, 7(3), 269–281. <https://doi.org/10.1080/14616730500245906>
- Slade, A., Aber, J. L., Bresgi, I., Berger, B., & Kaplan, M. (2004). *The parent development interview – revised*. Unpublished protocol. New York, NY: The City University of New York.
- Slade, A., Bernbach, E., Grienemberger, J., Levy, D., & Locker, A. (2004). *Addendum to fonagy, target, steele, & steele reflective functioning scoring manual for use with the parent development interview*. Unpublished Manuscript. New York: The City College; Graduate Center of the City University of New York.
- White, I. R., Royston, P., & Wood, A. M. (2011). Multiple imputation using chained equations: Issues and guidance for practice. *Stat Med*, 30(4), 377–399. <https://doi.org/10.1002/sim.4067>
- Zeegers, M. A. J., Colonesi, C., Stams, G.-J. J. M., & Meins, E. (2017). Mind matters: A meta-analysis on parental mentalization and sensitivity as predictors of infant-parent attachment. *Psychol Bull*, 143(12), 1245–1272. <https://doi.org/10.1037/bul0000114>
- Zigmond, A. S., & Snaith, R. P. (1983). The hospital anxiety and depression scale. *Acta Psychiatr Scand*, 67(6), 361–370. <https://doi.org/10.1111/j.1600-0447.1983.tb09716.x>
- Zimet, G. D., Dahlem, N. W., Zimet, S. G., & Farley, G. K. (1988). The multidimensional scale of perceived social support. *Journal of Personality Assessment*, 52(1), 30–41. https://doi.org/10.1207/s15327752jpa5201_2